The three-second "subjective present": a critical review and a new proposal

Abstract

It has been argued that there is a "subjective present" or "experienced moment" of about three seconds in duration, involving automatic binding of events into perceptual units on that time scale. Research on topics that have been taken as relevant to this proposal is reviewed. The topics include accuracy in reproduction of stimulus durations, synchronization of behaviour with a regular beat, mental rhythmization of a regular beat, time units in behaviour, segmentation of observed behaviour into meaningful units, time scale of reversals of perception with bistable ambiguous figures, time scale of inhibition of return in visual search, and EEG responses to deviant stimuli in series of repeating stimuli. Most of the research findings were not consistent with the three-second window hypothesis. The small amount of supportive evidence is better interpreted as effects of specific processing mechanisms, not as showing general temporal integration. The evidence shows that temporal integration occurs on multiple time scales and no particular duration is special, and that windows of temporal integration are defined in terms of information density, not in terms of duration. The subjective present is constructed through local temporal integration on multiple time scales, further integrated into a coherent global representation of what is going on.

Keywords: subjective present; experienced moment; three-second moment; temporal integration; temporal window

The three-second "subjective present": a critical review and a new proposal

The time span of perception covers more than just the moment that is present now. Imagine watching a ball that has been thrown through the air. A single present moment in the flight of the ball would be the equivalent of a short exposure photograph. Such a photograph shows where the ball is at a moment in time, but reveals nothing of its history. At each successive moment, the previous photograph would cease to exist and a new, equally momentary one would replace it. This is nothing like the way we perceive the world. Clearly the brain is pulling off a very special kind of trick, retaining information about the recent past and integrating it to construct perception of the world as not only spatially but also temporally coherent, comprising not just things, but things <u>happening</u>, moving, changing, continuing as they are, and doing things. Every time we hear what someone is saying to us, feel the texture of some fabric, experience ourselves manipulating a tool, or see a car being driven along the street, those percepts are only possible because the brain is adept at integrating information over a short period of time to make a perceptual construct of happening.

In general, temporal integration involves sampling information input over a period of a time and processing it so as to generate some sort of unified percept. Temporal integration can operate in many ways and on many time scales. However, a simple example would be processing of auditory input, specifically speech sounds, to generate a percept of a word. This involves integration of information over about 700 ms (depending on the duration of the spoken word; Stephens, Honey, & Hasson, 2013) and separation of the integrated information from its perceptual context, especially the speech sounds that precede and follow it. The resultant percept is unitary, coherent, and distinct from its context.

On what time scales does temporal integration occur? Certainly it occurs on the millisecond time scale and even less. For example, sound input to each ear can be integrated to form a percept of the spatial location of the sound source, with temporal differences on the microsecond scale (Grothe, 2003). Successive stimuli applied to the skin at intervals of 1 - 2 milliseconds can be temporally integrated to give rise to percepts of texture (Mackevicius, Best, Saal, & Bensmaia, 2012; Weber, Saal, Liber, Cheng, Manfredi, Dammann, & Bensmaia, 2013). However, it occurs on much longer time scales as well. For example, cross-modal information is integrated on a time scale of at least 200 ms to construct a synchronised percept of auditory and visual information (van Wassenhove, Grant, & Poeppel, 2007; Vroomen & Keetels, 2010). Successive visual stimuli can be integrated to form a combined, unitary percept when separated by as much as 240 ms (Akyürek & Wolff, 2016). Perception of the loudness of an auditory stimulus can involve temporal integration of information over several hundred milliseconds (Räsänen & Laine, 2013). Visual motion can be detected by integrating information across about 100 ms (Snowden & Braddick, 1990) but, under conditions of noise, integration can occur over as much as 3,000 ms (Burr & Santoro, 2001). Indeed, it has been argued that temporal integration in perception is a decision making process that continues until some decision criterion or signal detection threshold is reached, which means that it has no fixed time scale (Seilheimer, Rosenberg, & Angelaki, 2013).

The research just discussed might give the impression that the characteristics of temporal integration and the time scale on which it operates vary from one process to another, and from one occasion to another even when the same basic process is involved, at least in perceptual processing. However, it has been argued that there is a global form of temporal integration operating on a longer time scale, that essentially generates an integrated, coherent, overall sense of what is going on. Terms used for this overall impression include the "specious present" (James, 1890), the "subjective present" (Pöppel, 1997, 2009), and the "experienced moment" (Wittmann, 2011). I shall adopt the term "subjective present", but in quotation marks as a reminder that it is not every author's preferred term. The hypothesized time scale of integration has varied somewhat. James (1890) proposed a duration for the "specious present" of about 6 s. Fraisse (1984) proposed that the specious present has an upper limit of about 5 s but an average value of 2 - 3 s, thereby recognising that its duration is not absolutely fixed. More recent publications have settled on a duration around 2 - 3 s (Montemayor & Wittmann, 2014; Pöppel, 1997, 2009; Wittmann, 1999; 2009, 2011, 2013; Wittmann & Pöppel, 1999/2000). The proposal of global temporal integration to generate a "subjective present" on a time scale of a few seconds is the central focus of this paper.

The main body of this paper begins with a brief summary of the characterisation of the "subjective present" in the recent literature. Following that is a critical review of relevant research literature, that is, the research cited in the recent publications (Montemayor & Wittmann, 2014; Pöppel, 1997, 2009; Wittmann, 1999; 2009, 2011, 2013; Wittmann & Pöppel, 1999/2000) as support for the proposal. Two general points that emerge from the review can be highlighted here. One is that, collectively, the research findings do not converge on a particular duration or even a range of durations for the "subjective present". At the supra-second scale as well as the sub-second scale, temporal integration is a local phenomenon, operating in different ways in different processes. The second point is that much of the evidence that has been interpreted as supporting the hypothesis of a "subjective present" does not, in fact, appear to involve temporal integration, and therefore its relevance to the hypothesis of a "subjective present' is dubious.

After the literature review I put forward a new proposal as to what the "subjective present" might be, that takes these points into account. The three main features of this proposal are that temporal integration operates in different ways in different processes, that temporal parameters of the "subjective present" are set by information density rather than time per se, and that there is some form of global integration of the products of local integration processes that is responsible for the subjective coherence of experience and the general sense of ongoing happening.

The subjective present: a brief characterisation

The two authors who have written most on the "subjective present" in recent times have been Pöppel and Wittmann, and in this section I briefly summarise their characterisations of the "subjective present".[1]

Pöppel (1997, 2009) set the "subjective present" in the context of a hierarchical model of temporal perception comprising temporal integration windows with resolutions of 30 ms and 3 s, respectively. Pöppel (2009) distinguished between "two mechanisms, one being pre-semantic providing a temporal frame for processing and the other being responsible for the content of what is processed" (p. 1888). Thus, the pre-semantic mechanism establishes a temporal frame of about 3 s and perceptual information is then integrated within that frame. He proposed that the binding is not governed by the specifics of processing but by a general "neuro-cognitive machinery" (p. 59) operating with intrinsic time constants. As a result, the 3 s window of integration is a general feature that should be observed in many different kinds of processes. Pöppel (2009) added that "conscious activities are temporally segmented into intervals of a few seconds" (p. 1891), thereby associating the "subjective present" with consciousness, and suggesting that the "subjective present" proceeds in discrete windows of approximately 3 s duration. Functionally, the temporal window supports a complementarity of identity and dynamics: "the brain creates temporal windows of just a few seconds within which the identity of a percept or a concept is maintained (stationarity), and allows after such an interval the access of a new percept or concept (dynamics)" (p. 1893). Pöppel did not say what temporal integration is, only that it is pre-semantic and automatic; nor did he propose a specific mechanism for the kind of temporal integration that would generate a "subjective present". He did say that the answer to the question of how long the "subjective present" lasts "can be derived from a number of different experiments, all of which converge to a value of approximately 2-3 s" (Pöppel, 2009, p. 1891). This is a striking claim but, as we shall see, it is incorrect.

The "subjective present" has also been discussed by Wittmann (1999, 2009, 2011, 2013), who preferred the term "experienced moment"; here I shall focus on the most recent version of Wittmann's account, by Montemayor and Wittmann (2014). Montemayor and Wittmann postulated three levels of the present, one corresponding to the experience of simultaneity, the second characterised by the experience of duration, and the third characterised by narrative autobiography. The second of these corresponds to the "subjective present" or "experienced moment" (e.g. p. 331). They proposed that "[t]he functional principle of this type of experienced present is that once a set of contents is unified into a single experience, such experience appears subjectively as a unique conscious episode, continuously interrelated with other experiences" (p. 331). Like Pöppel, Wittmann (1999, 2009, 2011) used the term "temporal integration" frequently, but did not define it nor propose a specific mechanism for the kind of integration that would generate a "subjective present". In particular, Montemayor and Wittmann (2014) did not say what the difference is between the temporal integration that results in a single experienced moment and the continuous interrelation that appears to link different experienced moments.

Both Pöppel (2009) and Montemayor and Wittmann (2014) stressed that the hypothesized three-second window should not be taken too narrowly. Montemayor and Wittmann proposed a lower boundary for the "subjective present" of 300 to 500 ms and an upper boundary of about 3 s. Thus, any kind of integration on a time scale between 300 ms and 3 s could be taken as evidence for the proposed experienced moment. Pöppel (2009, p. 1891) wrote that the subjective present "does not have the typical characteristics of a physical constant with precise values, but it represents an operating range of a few seconds; in addition, one has to expect some interindividual variability for such a biological process". Because of this, it is difficult to specify precise falsification criteria for the proposal. The idea of a "subjective present" or "experienced moment" implies something that unifies all of experience, or something that occurs over the whole of experience. Thus, even if there is a large range of durations that would be compatible with the hypothesis, there should be some consistency in mean values across different areas of experience or different kinds of processing.

Several distinct lines of research have been taken by Pöppel (1997, 2009) and Wittmann (Montemayor & Wittmann, 2014; Wittmann, 1999, 2009, 2011) as supporting the hypothesis. Each of these will be reviewed below. The focus will be not just on scrutinising and evaluating the evidence for mean values within each line of research: there are important issues concerning whether the time scales are similar across different processing domains, whether an identifiable temporal integration mechanism occurs or not, whether it is automatic, as it should be under the hypothesis of the "subjective present", and whether there are similar and consistent mechanisms across processing domains. The authors did not make a case for the relevance to the hypothesized "subjective present" of any of the research they reviewed: this is a critical issue that will also be addressed. Where I refer to the 3 s time scale, this should be understood as allowing some temporal leeway in either direction.

Is the "subjective present" concerned with temporal units of consciousness, or temporal units of perception? The authors of the proposals referred to consciousness (see quotations above), but the products of many kinds of processes, both perceptual and post-perceptual, may form part of the contents of consciousness, so consciousness, whatever that may be, should not be taken as defining the "subjective present". The processes of temporal integration that generate the "subjective present", as hypothesized by the authors of the proposals, go beyond the likely temporal limits of perceptual processing, but take the products of perceptual processing as their input. Whether their output should be regarded as a perceptual product is hard to judge. If iconic store, which operates on the sub-second time scale (Sperling, 1960), is a form of memory, then the "subjective present" should perhaps be regarded as a form of memory as well; but both of those things also seem to concern what is going on now, so they are more closely tied to perception than, for example, information in working memory is.

Review of research

Several lines of research have been interpreted by Pöppel (1997, 2009) and Wittmann (Montemayor & Wittmann, 2014; Wittmann, 1999, 2009, 2011) as evidence supporting the proposed 3 s "subjective present". The present review assesses the evidence from those lines of research and is specifically concerned with the issue of the "subjective present": other issues, such as hypothesized processes specific to one line of research, are only considered if they have direct relevance to the issue of the "subjective present". The most comprehensive of the previous reviews is that by Pöppel (2009) and the research he cited will be evaluated. However, many relevant studies were not mentioned in any of the previous papers, and consideration of these permits a more thorough assessment of the case for the hypothesized "subjective present".

Duration reproduction

In studies of duration reproduction, stimuli of set durations are presented and participants then attempt to reproduce the duration, for example by pressing a button for what they judged to be the correct amount of time. Pöppel (1997, 2009) cited some studies as supporting the hypothesis of a 3 s "subjective present" on the grounds that they found evidence for different processing modes for reproducing durations below and above 3 s (Elbert, Ulrich, Rockstroh, & Lutzenberger, 1991; Kagerer, Wittmann, Szelag, & von

Steinbüchel, 2002; Szelag, Kowalska, Galkowski, & Pöppel, 2004). Several other studies have yielded relevant evidence and will be considered here.

The studies cited by Pöppel (1997, 2009), to which can be added Szelag, Kowalska, Rymarczyk, and Pöppel (2002), presented similar sets of durations, a large set with 0.5 s increments from 1.0 s to 5.5 s, and a small set with 0.5 s increments from 1.0 s to 3 s. In all the studies there was a tendency for underestimation to occur increasingly as duration increased (except for children with autism in the study by Szelag et al., 2004). The authors argued that the data were best fitted by two regression lines with different slopes either side of 3 s, indicating accurate judgment at durations below 3 s and underestimation at durations above 3 s. In fact this is not clear in the data: in all the studies, the tendency to underestimate duration increased steadily as duration increased, and there is no clear evidence of a discontinuity in the trend around 3 s. The results appear consistent with a gradual increase in error, possibly with a curvilinear function. There is no clear evidence that this would fit the data less well than the two straight regression lines proposed by the authors.

Two studies on duration reproduction have found evidence suggestive of two different ways of reproducing durations with a temporal boundary between 2 and 3 s. Woodrow (1930) studied reproduction of durations ranging from 200 ms to 30,000 ms. He measured variability in duration reproduction as a function of the mean reproduced duration at each duration value. The mean variability was approximately stable at around 8 - 10% for durations from 200 to 2,000 ms, and abruptly jumped to about 16 - 17% for longer durations. Participants reported that reproduction of these longer durations felt like guesswork, but that of the shorter durations did not.

In a study by Ulbrich, Churan, Fink, & Wittmann (2007), participants performed a duration reproduction task under cognitive load. Visual and auditory stimuli were used. The basic reasoning was that temporal integration processes on the scale of the "subjective present" are automatic, as Pöppel (1997, 2009) hypothesized. In that case, duration reproduction for durations less than 3 s should not be affected by cognitive load, whereas durations longer than 3 s, which are estimated by some more attentive, controlled process, should be affected by it. The stimuli they used were 1000, 2000, 3000, 4000, and 5000 ms in duration. Two features of the results supported the reasoning. They found that individuals with better working memory were more accurate than others at 4000 and 5000 ms for auditory stimuli and at 3000 and 5000 ms for visual stimuli, but not at the shorter durations. This supports the claim that durations around 3000 ms and above are estimated by a process that is not automatic, on the grounds that the performance of individuals with poorer working memory was more adversely affected by cognitive load at those longer intervals. The lack of significant difference at intervals less than 3000 ms is consistent with the hypothesis that the duration reproduction mechanism at that time scale is automatic. A factor analysis showed that short intervals loaded on one factor for both modalities and long intervals loaded on two modality-specific factors. The implication is that short intervals were judged by an automatic process and long intervals by a non-automatic process, supporting the hypothesis.

The use of a concurrent task that imposes a cognitive load is critical for this research. Without it, participants might use a counting or rhythmical movement strategy to estimate longer intervals: such methods in effect subdivide the interval to be reproduced into a series of short intervals, each of which can be generated accurately, so that accurate judgment can be generated across a wide range of intervals by counting the short intervals. This rules some otherwise relevant studies out of contention, such as Gibbons and Rammsayer (2004). Woodrow (1930) did not use a cognitive load task, but he instructed participants not to count or make any rhythmical movements, and not to attend to their breathing. He also asked the participants to report on what they had been doing during the task and, as far as he could ascertain, they complied with the instructions.

Other studies that used cognitive load tasks have found different results. Fortin and Couture (2002) used a range of durations between 1.85 and 6.45 s and found that the effect of increasing cognitive load was similar at all durations, which does not support the hypothesis of a change from automatic to controlled processing between 2 and 3 s. Macar, Grondin, and Casini (1994) found effects of dividing attention between two tasks that were similar for stimuli with durations of 1.5 and 3 s. Brown (1997) also found similar effects of increasing cognitive load with durations of 2 s and 5 s.

Rammsayer and Lima (1991) found an effect of cognitive load for duration discrimination with durations longer than 1000 ms but not for one with discriminations in the millisecond range. Duration discrimination (deciding which of two intervals is longer) is not quite the same as duration reproduction, but both depend on timing mechanisms and on short-term memory, because stimulus and reproduction, in one case, and two stimuli, in the other, are sequential and judgment or reproduction must both be, to some degree, retrospective. Rammsayer and Ulrich (2011) considered a model in which there is what they called a sensory (presumably automatic) mechanism for processing durations up to about 500 ms and a cognitively controlled one, that is affected by cognitive load manipulations, for durations longer than that. They found support for the model in their research. This would fit with the results of the studies by Fortin and Couture (2002), Macar et al. (1994), and Brown (1997), in which only durations longer than 500 ms were used.

Separate processing mechanisms for durations below and above 500 ms does not rule out the possibility of separate mechanisms below and above 3 s, though it is not predicted by the hypothesized "subjective present", either in Pöppel's (1997, 2009) version where the "subjective present" is 2 - 3 s in duration or in the version by Montemayor and Wittmann (2014) where the lower bound is about 300 ms. However the evidence for effects of cognitive load at durations between 500 ms and 3 s is quite damaging to the hypothesis of a 3 s "subjective present", because temporal integration is supposed to be automatic on that time scale. Moreover, the evidence for different mechanisms below and above 3 s is not compelling: there is no sign of it in the cognitive load studies by Fortin and Couture (2002), Macar et al. (1994), and Brown (1997), and the evidence for a difference in processing reported by Kagerer et al. (2002) and Szelag et al. (2002, 2004) is not strong enough. There is always the possibility that results may be affected by methodological factors that remain to be ascertained, but that possibility applies to the studies by Woodrow (1930) and Ulbrich et al. (2007) as well as to the other studies reviewed here. Without direct testing it is impossible to know how results might be affected by features of the specific methods used.

The studies reviewed here constitute a small portion of a voluminous literature on time perception and time estimation (see, for example, Block & Grondin, 2014; Matthews & Meck, 2016). If there is a critical duration in time perception and estimation, it lies between 1 s and 1.5 s. Many studies have shown that error and variability in time estimation or reproduction increase in a way that violates Weber's law beyond that approximate duration, indicating an additional source of error in timing at those longer durations (Gibbon, Malapani, Dale, & Gallistel, 1997; Grondin, Laflamme, & Mioni, 2015). Further discussion of this extensive literature and possible explanations for the results lies outside the scope of this review. However, two points of relevance need to be made. One is that issues of time perception and subjective duration are primarily addressed by hypotheses of timing mechanisms. For example, the best known hypothesized timing mechanism, the pacemaker/accumulator model (Treisman, 1963), involves a pacemaker that emits pulses at regular intervals, which are collected and stored in an accumulator, from which they can be retrieved to generate a judgment of the duration between two events. This is not temporal integration, and it is certainly a different kind of process. It is questionable, therefore, whether the evidence from duration reproduction and estimation research is even relevant to the issue of temporal integration in the construction of a "subjective present". The other issue is that it is not clear whether the errors observed in duration reproduction occur at encoding, in retention, or in reproduction. A duration might be inaccurately encoded or it might be accurately encoded but inaccurately reproduced. It is doubtful whether the integration time of perceptual processing for the "subjective present" can be ascertained under those circumstances.

Sensorimotor synchronization

In this line of research a regularly repeating stimulus, such as a metronome beat, is presented, and participants are instructed to tap in synchrony with the stimulus. It has been claimed that performance on this task varies qualitatively depending on the temporal interval between successive beats. With intervals below 3 s, taps are relatively accurate and tend to be anticipatory, occurring slightly before the beat. With intervals above 3 s, timing of taps is more irregular and reactive, as if the participants are no longer able to predict exactly when the beat will occur (Pöppel, 2009). Pöppel (2009) and Tanida and Pöppel (2006) argued that this reflects a form of temporal integration that affects goal-directed behaviour. Tanida and Pöppel (2006) proposed a model of goal-directed behaviour with five levels of temporal representation. The 3 s temporal integration window is the second of these and it represents functions associated with anticipatory control such as (in their example of driving a car), the basic behavioural operations of steering, braking, and so on. In general, anticipatory control is claimed to be another instance of temporal integration on the 3 s time scale. In this section, I review the evidence concerning the hypothesized qualitative shift in sensorimotor synchronization around the 3 s interval.

There is evidence for a qualitative change in the behavioural control mechanism around stimulus onset asynchronies (SOA) of 300 ms (Peters, 1989; Wing & Kristofferson, 1973). This is shown by a sudden increase in variability of inter-tap intervals around SOA = 300 ms. Peters (1989) argued that the increase in variability marked a shift from automatic to controlled movement, but the exact mechanisms involved were not specified. One possibility is that, with SOA < 300 ms, taps are produced at a regular rate, and the rate is synchronized by a mechanism that utilises information about the rate of stimulus beats, perhaps integrated over several beats. With SOA > 300 ms, there is time for an error correction process that is guided by the assessed discrepancy between the previous beat and the previous tap (see, e.g., Blakemore, Frith, & Wolpert, 1999; Blakemore, Wolpert, & Frith, 2002; Flanagan & Wing, 1997; Flanagan & Johansson, 2003; Jacoby & Repp, 2012; Repp & Su, 2013). The error correction system utilises the detected error (the temporal difference between the auditory perception of the metronome beat and the internal model of the motor output and its timing) to guide the timing of the next tap. The error correction process continues in that way from tap to tap.

The question is whether that error correction mechanism reaches some kind of temporal limit around SOA = 3 s. There is evidence consistent with that hypothesis. Mates, Radil, Müller, and Pöppel (1994) presented tones with SOA varying from 300 ms to 4800 ms. Up to SOA = 1800 ms there was a tendency for taps to occur slightly ahead of the beat. At SOA > 1800 ms two different tendencies were apparent. One was a noticeable increase in variability of timing around the time of the beat, as if participants were still trying to predict the occurrence of the beat but were unable to do so accurately. The other was a tendency for taps to occur after the beat, with a latency suggestive of reacting to the beat and not predicting it.

What seems to have been the earliest study of sensorimotor synchronization, by Woodrow (1932), yielded evidence partly consistent with the results reported by Mates et al. (1994). Woodrow tested SOAs in the range 250 to 4000 ms and found a consistent anticipatory tendency at all SOAs, which does not agree with the change to reactive responding reported by Mates et al. However he also noted a qualitative difference between tapping at intervals of 2000 ms or greater and at shorter intervals: participants reported that at shorter intervals they could proceed automatically or rhythmically, whereas at longer rates they described having to treat each tap as an individual event. On the basis of data on errors of synchronization, Woodrow estimated the upper limit of the automatic or rhythmic synchronization to be at 3.34 s, which is consistent with the proposal of temporal integration on a time scale of about 3 s.

Repp and Doggett (2007) found that timing variability increased linearly with SOA up to 2500 ms, and then showed a steeper increase, consistent with the findings reported by Mates et al. (1994) and suggestive of a qualitative change. However, when participants were instructed to predict the next beat and not to react to it, they were able to do that and the linear increase in variability continued without any indication of discontinuity. This suggests that a voluntary switch to reactive tapping occurs when it becomes easier than predicting, not because there is a mandatory switch from one behaviour control mechanism to another.

Miyake, Onishi, and Pöppel (2004) used a cognitive load task and found that this affected tapping at SOA = 1500 ms and above, by significantly reducing the proportion of anticipatory tapping. The effect of cognitive load disappeared at the longest SOA, 6000 ms, but this is probably because tapping was mostly reactive at that interval. This supports the suggestion that participants switch to reactive tapping when the attempt to synchronize accurately becomes more cognitively demanding and no more accurate than reactive tapping. If temporal integration is automatic, as is hypothesized in Pöppel's (1997, 2009) account of the "subjective present", then the effects of cognitive load in this research show an upper limit on the "subjective present" of about 1500 ms.

Other studies have not supported the hypothesis of a qualitative change in performance with SOA around 2 - 3 s. Matsuda, Matsumoto, Furubayashi, Hanajima, Tsuji, Ugawa, and Terao (2015) used SOAs between 200 and 4800 ms, and their participants were patients with cerebellar ataxia and healthy controls. Both patients and controls exhibited accurate inter-tap intervals across the whole range of SOAs except that the ataxia patients had problems with the shortest intervals, 200 - 333 ms. Synchronization error gradually increased as SOA increased but there was no indication of a change in the rate of increase at any point. Synchronization error increased more rapidly for the ataxia patients than for the controls but reached a peak at 1800 ms and then levelled off.

There are indications that the tendencies reported in the research may be specific to the tapping task. There are other ways of testing sensorimotor synchronization. Van der Wel, Sternad, and Rosenbaum (2010) asked participants to move a dowel back and forth between two target locations in time with a metronome beat, with intervals ranging from 370 ms to 1667 ms. Tapping the targets tended to occur ahead of the beat by about 20 ms, but this was approximately constant across the whole range of durations and did not diminish. The longest duration is still shorter than that at which reactive tapping tends to take over, but the results differed from those of standard tapping tasks in that anticipatory tapping did not diminish and error did not increase with increasing SOA. The movement of the dowel between two target locations makes the behaviour effectively continuous at all intervals, whereas a simple tap becomes discontinuous when SOA exceeds the time taken to raise and lower the finger. This could be an important factor in determining the tendencies in timing that occur.

Takano and Miyake (2007) distinguished between phase correction (controlling the coincidence of tap timing with stimulus timing) and period correction (synchronising tap period with SOA). They reported evidence for two phase correction mechanisms. One was a negative feedback mechanism, perhaps the error correction mechanism discussed earlier, and was found at all values of SOA in their study (450 to 1800 ms). The other was found only at intervals between 1200 and 1800 ms, and was disrupted by a cognitive load task. They suggested that it could represent the operation of a cognitive timing mechanism for anticipation of the next beat.

Overall, then, there are indications of changes of behaviour control mechanism in tapping around SOA = 300 ms (Peters, 1989; Wing & Kristofferson, 1973), 1200 ms (Takano & Miyake, 2007), 1500 ms (Miyake et al., 2004), 2500 ms (Repp & Doggett, 2007), and approximately 2400 to 3600 ms (Mates et al., 1994). Other studies have found no evidence of qualitative change in behaviour control with increasing SOA (Matsuda et al., 2015; van der Wel et al., 2010). The tendency to switch from anticipatory to reactive tapping at longer values of SOA may indicate a trade-off between accuracy and difficulty, with a voluntary switch to the less cognitively demanding reactive tapping when timing error is not significantly greater with reactive than with anticipatory timing. Since the evidence for a qualitative change around 2 - 3 s is mainly that switch from anticipatory to reactive timing (Mates et al., 1994), the finding that reactive timing can be abolished by a simple instruction to predict on all trials (Repp & Doggett, 2007) considerably undermines the case for a temporal integration window of 2 - 3 s.

It is clear that there is much still to learn about the mechanisms operative in sensorimotor synchronization, especially concerning the extent to timing phenomena depend on whether behaviour itself is continuous (van der Wel et al., 2010) or discontinuous (see also Cos, Girard, & Guignon, 2015). In the present state of knowledge, the small amount of evidence consistent with a temporal integration window of 2 - 3 s is outweighed by evidence for qualitative changes at other intervals, evidence against qualitative changes in behaviour patterns, and the case that the supposed evidence for the window actually indicates a voluntary change of strategy. However, as with duration reproduction, there is no compelling case that temporal integration is involved in sensorimotor synchronization. The basis for synchronization would appear to be a timing mechanism, coupled with motor co-ordination. Changes of behaviour control mechanism as SOA increases may indicate changes from one form of timing to another, and control of behaviour may become less automatic and more controlled as SOA increases, but neither kind of change need involve temporal integration. Timing, as has already been said, can be governed by a mechanism that counts pulses issued at regular intervals (Treisman, 1963). If that can be described as temporal integration, it is of a very specific kind, and there is no reason why its operating characteristics should apply to other kinds of temporal integration. The relevance of this research to the issue of the "subjective present", therefore, remains to be established.

Subjective rhythmization

In subjective rhythmization, a steady beat is presented, e.g. by metronome, and participants are asked to accentuate beats mentally at regular intervals to create a subjective rhythm. Pöppel (2009) argued that studies of this by Szelag and colleagues showed that this becomes impossible if the interval between beats exceeds a few seconds, and that this was further evidence for a boundary on temporal integration of approximately 3 s.

In the first study, by Szelag, von Steinbüchel, Reiser, de Langen, and Pöppel (1996), beats were presented at rates between 1 and 5 per second. Participants were asked to integrate the beats into larger units of 2, 3 or more by mentally accentuating beats to create a subjective rhythm. The participants reported how many beats they could accommodate into a single unit of rhythm at the different presentation rates. The authors also compared different age groups of younger and older adults. Overall, the mean integration interval (the number of reported beats used as the basis for the subjective rhythm multiplied by the temporal difference between them) was 1351 ms for the younger participants and 1751 ms for the older ones. This varied with beat frequency, however, and became shorter as frequency increased, from about 2900 ms at 1 beat per s to 1028 ms at 5 beats per s. There was considerable variance: for example, at 1 beat per s, individual integration intervals ranged from about 1300 ms to about 4400 ms. When beats were separated by more than 1 s, "an accentuated rhythm could no longer be perceived and only isolated beats were heard" (p. 222).

Szelag, von Steinbüchel, and Pöppel (1997) repeated the experiment with patients with various kinds of brain damage and healthy controls. They presented their results as a graph plotting integration interval against beat frequency. On the hypothesis of a 3 s temporal integration interval, the results should approximate to a straight line across the number of beats per second at the 3000 ms level. This was not the case, however. For the controls and all categories of patient other than those with Broca's aphasia, the integration interval peaked at about 2400 ms at 1 beat per s and declined steadily to about 1100 ms at 5 beats per s. This trend closely resembles that found by Szelag et al. (1996). Other studies have found similar results. Bolton (1894) found the outer limit of intervals between beats for subjective rhythmization to be 1580 ms with variation of about 290 ms. Baath (2015) found that, with SOA of 1500 ms or more, subjective rhythmization occurred on less than half of the trials, but that there was large inter-individual variability.

Although the number of studies is not large, there is an impressive consistency in the findings across more than a century of research. The limit on subjective rhythmization is an SOA about 1500 ms, but with substantial inter-individual variability and also varying with beat frequency. This is shorter than the 2 - 3 second window usually taken as delimiting the "subjective present" (Pöppel, 2009; Wittmann, 2011). The main issue concerns whether subjective rhythmization reflects a general temporal integration on a time scale of a few seconds, or something specific to auditory information processing. Baath (2015) reviewed two classes of explanations for subjective rhythmization, both of which are specific to the auditory perception of events with regular periodicity. To illustrate, one class of explanation is resonance theory. According to this account, a regular repeating sound entrains neural oscillatory circuits, selectively increasing the amplitude of those that have a similar oscillatory period to the period of the beat. This generates the experience of beat and forms the substrate for rhythmization. There are several variants on this basic account (Baath, 2015). Baath's study yielded evidence supportive of the resonance theory account, so it is probably the front runner in the field at present. Presumably the upper limit on beat interval for subjective rhythmization is set by the longest period of an oscillatory circuit in the auditory system. This is not a temporal integration mechanism. Therefore, if this explanation is correct, the fact that the upper limit on beat interval for subjective rhythmization is not far from the proposed temporal integration interval for the "subjective present" would be a coincidence. On the other hand, perhaps it could be argued that the upper limit on the time scale of neural oscillations is a general feature of brain activity, both defining and explaining the "subjective present".

Time units in speech and other behaviour

Pöppel (2009) argued that the duration of intentional movements tends to be about 2 - 3 s, citing as supporting evidence two studies of field observations of human behaviour (Schleidt, Eibl-Eibesfeldt, & Pöppel, 1987; Schleidt & Kien, 1997). In fact there is a substantial literature on this. I shall briefly summarise the main findings and then comment on some methodological features of the research.

A series of studies by Schleidt and colleagues reported analyses of behaviour durations in videotaped recordings of spontaneous behaviour in several cultures, including Europeans, Trobriand Islanders, Yanomami (Amazonian) Indians, and Kalahari Bushmen (Feldhütter, Schleidt, & Eibl-Eibesfeldt, 1990; Lemke & Schleidt, 1999; Schleidt, 1988; Schleidt et al., 1987; Schleidt & Feldhütter, 1989; Schleidt & Kien, 1997). Schleidt and Kien (1997) provided a table of kinds of behaviour included in the analyses, the most common of which were chopping or felling, peeling (e.g. stripping bark from a tree), transferring (e.g. pouring liquid from one pot to another), and wiping or rubbing. Some of these behaviours are repetitive, and the researchers used a criterion of at least three repetitions for a repetitive behaviour unit. The results, summed across behaviour types, show modal behaviour durations between 2 and 3 s in all the cultures studied. There was, however, a wide range of variation with durations as long as 9 s recorded in some instances (Feldhütter et al., 1990;

Schleidt et al., 1987). The greatest variation in duration in this set of studies occurred in the study by Lemke and Schleidt (1999) who focussed on behaviour units with the arms only. They found median durations of 1.8 s for non-repetitive behaviours and 2.8 s for repetitive behaviours; a similar difference between repetitive and non-repetitive behaviours was found by Schleidt and Feldhutter (1989). The variance was considerable, however, with some behaviour durations as long as 11 s.

An average (mean, median, or mode) behaviour duration of 2 - 3 s has been found in other studies as well. Nagy (2011) analysed the duration of spontaneous embraces between two people in footage from the Beijing Olympics and found a mean of 3.17 s. This varied somewhat depending on the relationship between the embracers: competitor/competitor embraces were shorter (mean = 1.81 s) than competitor/coach embraces (mean = 3.77 s). There were also cultural variations, but not outside the range 2 - 4 s. Turner and Pöppel (1983) analysed poetry reading in several languages and reported that reading a behavioural unit (a line) almost always took between 2 and 4 s. Gerstner and Cianfarani (1998) covertly filmed people eating and analysed the duration of chewing bursts. They found a mean duration of 2.91 s, but the mode was about 1 - 1.5 s, indicating a skewed distribution, and burst durations up to 16.5 s were observed.

There are, however, reasons for doubting the validity and generalisability of the findings reported so far. Po, Kieser, Gallo, Tésenyi, Herbison, and Farella (2011) also ran a study of chewing burst duration and found a mean duration of 13.0 s with 95% confidence intervals of 2.7 s and 34.9 s. The mean is clearly well beyond the 3 s "subjective present". The difference between these results and those obtained by Gerstner and Cianfarani (1998) can be attributed to choice of criterion for identifying the boundary of a chewing burst. Gerstner and Cianfarani defined a pause as 1.5 s or longer, but Po et al. defined it as 2 s or longer. Thus, under their criterion, two chewing episodes separated by a pause of 1.6 s would be recorded as a single burst, whereas in the study by Gerstner and Cianfarani they would be recorded as two bursts.

This raises perhaps the most important methodological issue for this line of research: how is a behavioural unit identified? There must be some objective criterion for identifying onset and offset of a behavioural unit, or for identifying pauses in an ongoing stream of behaviour. As the two chewing burst studies show, the mean duration of behavioural units depends critically on the criterion pause duration, and on any other choice of cue for identifying onset and offset of behaviours. This is most clearly illustrated in a study by Kien and Kemp (1997). They analysed phrase durations in spoken German and Korean, including both free speech and reading out written prose. Their criterion for identifying phrases was pause duration, but they used three different criterion values, 248 ms, 496 ms, and 1000 ms. Note that all of these are shorter than the criteria used in the chewing studies. Phrase duration varied substantially depending on the criterion chosen. For prose reading, with the 248 ms criterion, the modal phrase duration was < 1 s and the distribution tailed off rapidly to a limit at about 6 s. With 496 ms, the distribution was approximately flat from 0.5 to 5 s, with a smaller number of phrases up to about 10 s. With 1000 ms, phrase lengths were not normally distributed and there were as many in the range 8 - 10 s as in the range 0 - 2 s. In fact, there was an apparent dip in the distribution around 2 - 3 s, and some phrases were as long as 20 - 30 s. Similar patterns were observed in the other kinds of speech studied.

Returning to the set of studies by Schleidt and colleagues, I can find no information about pause duration criteria. Schleidt and Kien (1997), which is a review and discussion of that set of studies, reported criteria for identifying behavioural units which were apparently used in the whole set of studies, but they are not well justified. For example, "[a] sequence of movements was regarded as an action unit only if its beginning and end could be clearly seen" (p. 79). What counts as being "clearly seen" is not specified. "Two consecutive movements do not belong to the same action unit if they are not functionally related" (p. 79).

They did not say how the researchers decided whether two movements were functionally related or not. "A sequence in which the same movement pattern is repeated several times rhythmically is considered a single action unit if the movement patterns are morphologically identical, and if they succeed one another without any interruption or pause" (p. 79). This seems to indicate that any detectable pause counts as a boundary between action units, but perhaps only for repetitive behaviours. And the criterion duration of a pause is not specified. Overall, there is too much leeway for subjective judgment of boundaries. Of equal concern, some behaviours were excluded from the studies. Schleidt et al. (1987) stated that the repetitive behaviours they studied did not include working, eating, body scratching, games with rules, and ritual dances "because of probable outside influences" (p. 127). The authors did not explain what these outside influences might be, how they were thought to have affected behaviour units, or indeed why the behaviours <u>included</u> were deemed <u>not</u> to be subject to "outside influences". The effects of different pause duration criteria reviewed above clearly demonstrate that behaviour unit duration depends critically on how behaviour units are identified, not to mention on decisions to include or exclude various categories of behaviour. It is also necessary to point out that the researchers, who chose which kinds of behaviour to include, analysed the recordings and identified the behaviour units, were not blind to the hypotheses of the study.

I now focus specifically on speech. Speech comprises a set of hierarchical levels of temporal organisation. Those that might count as behaviour units include phonemes or syllables, words, phrases, sentences or similar grammatical units, paragraphs, and utterances. To illustrate, Vollrath, Kazenwadel, and Krüger (1992) analysed speech at two levels: "vocalization level where phonations (roughly corresponding to vowel-consonant units) are integrated, forming syllables and words (the vocalizations), and at the utterance level, where intonational phrases are combined, resulting in utterances" (p. 479). For vocalisations, they found a mean duration of 0.32 s for singular units and 0.69 s for combined ones. For utterances, they found a mean duration of 2.6 s for singular speech units and 4.5 for combined ones. They did not report how boundaries between utterances were established, nor how they identified combined utterances. However, the research suffices to make the point that there are multiple temporal levels on which speech could be analysed, and indeed it would be quite surprising if none of them was anywhere near 2 - 3 s.

Having said that, there is reason to think that there is something special about the 2 - 3 s duration in speech comprehension. Sachs (1967) ran a study in which participants listened to a passage of text being read. At various points a bell sounded and then a sentence from earlier would be repeated, either identical or with change to either grammatical form or semantics. There were three levels of delay, zero, 80 syllables (about 27 s) and 160 syllables (46 s). At zero delay both kinds of change were identified with high levels of accuracy. At 160 syllable delay, semantic changes were still identified correctly on 78% of occasions, whereas correct identification of grammatical changes was about 60 % for passive/active changes and 52% for formal changes (where 50% is chance responding). Sachs (1974) used four values of delay, 3, 7.5, 12.5, and 23 s and found little decline in memory for semantic changes, but substantial declines for various kinds of formal changes even after short delays.

The studies by Sachs are an initial indication that semantic information is retained better on a time scale of seconds than syntactic information is, but have not yet pinned down 3 s as a critical duration for this difference. Roll, Lindgren, Alter, and Horne (2012) pointed out that studies of short-term memory for word lists show that substantial decay occurs in the first 2 - 3 s after presentation if rehearsal is not possible. For example, Peterson and Peterson (1959) found a decay curve with a negative exponential reaching asymptote after about 18 s, but with half of the decay occurring in the first 3 s. Roll et al. argued that this would lead to a tendency to insert prosodic boundaries when phonological traces start decaying, and so, regardless of the number of clauses, a prosodic phrase would tend to occupy a duration of 2 - 3 s. Presentation rate is the key to testing this. If the number of clauses presented per unit time is what matters, the integration boundary should be defined in terms of number of clauses, and could vary a lot in duration. But if it is defined by a temporal integration/retention window of 3 s, then it should not vary with time, no matter how many clauses are presented per unit time. Roll et al. looked at presentation rates in which three clauses took either 2.7, 4.0, or 8.1 s to present. ERP data supported the hypothesis that a prosodic boundary was inserted at intervals of about 2.7 s, regardless of speed of clause presentation. This supports the hypothesis of a temporal retention window of about 3 s.

Roll, Gosselke, Lindgren, and Horne (2013) investigated grammatical agreement, for example (in English) that "knows" agrees with "he" as the subject of the verb, but "know" agrees with "they" as the subject. They suggested that, when reading a word, the word is a cue to search short-term memory for grammatical forms that agree with it. A target word can be found if it was presented less than 3 s ago, but at longer times ago it is less likely to be found and the reader has to search outside the 3 s window, presumably for a semantic representation. They looked for brain responses that would indicate the occurrence of this kind of switch in search, by varying presentation rate to create different temporal gaps between the agreeing or disagreeing words. They reported evidence for a switch from one kind of search to another between 2.5 and 3 s temporal gaps, which is again consistent with the hypothesized temporal retention window of about 3 s.

The studies by Roll et al. (2012, 2013) both show that a processing bottleneck caused by rapid decay of information in short-term memory generates phenomena - prosodic boundary insertion and change from one kind of processing to another - on the time scale of 2.5 - 3 s, which coincides with that proposed for the "subjective present". Is this more than a coincidence? The involvement of the short-term memory decay curve suggests that it might be, although the decay curve is for verbal materials and it is not clear that it would generalise to other forms of information, such as visual stimuli (e.g. Luck & Vogel, 2013). If some kinds of information, such as the semantic content of utterances, preferentially survive the bottleneck, then the "subjective present" could mark the temporal limit on integration of kinds of information that do not get through the bottleneck. The fact remains, however, that behaviour durations depend on the criteria, such as pause length, chosen to divide behaviour into units, and for that reason alone the behaviour unit studies cannot be regarded as supporting the hypothesized 3 s "subjective present". One further caveat is that the bottleneck functions only when rehearsal is not possible, that is, when there is no spare cognitive capacity available, and it is not clear to what extent that is the case in social interaction and discourse. On the other hand, the very need for rehearsal to retain information that would otherwise decay suggests a transition to a different kind of processing. I shall return to this issue after completing the literature review.

One further observation about speech is pertinent. Nobody in the research reviewed in this section mentioned breathing. The normal respiration rate in adults is about 12 - 16 breaths per minute, which gives a duration of about 3.75 - 5 s per breath. This is not far off the mean phrase length found in reading prose, and certainly close to the mean of 4.5 s for combined utterances reported by Vollrath et al. (1992). However, Grosjean, Grosjean, and Lane (1979) found that, when reading out written sentences, identical pause structures occurred regardless of whether the reading used one breath or several. Pauses, and pause durations, are closely related to syntactic structure (Grosjean et al., 1979; Krivokapić, 2007), but not to breathing, so the hypothesis that breathing determines behaviour unit duration in speech can be ruled out.

Finally in this section, it is again pertinent to consider whether temporal integration is involved in regulation of behaviour units at all. If a temporal boundary in behaviour is attributable to decay of information in short-term or working memory, that is a phenomenon of loss of information and the need to compensate for it, not a phenomenon of temporal integration. Other temporal factors in behaviour are governed by practical considerations such as, in the case of chewing, the toughness of the food being chewed or, in the case of peeling bark from a tree, the point at which the bark happens to break. It is hard to see how temporal integration is involved in setting the time scale of such things.

Event segmentation

Event segmentation concerns how people perceive or construct temporal segments in social interaction or behaviour sequences. Such activities tend to be segmented at salient boundaries, so studying how people make segmentations can help to reveal their comprehension of events and event sequences in social interaction (Zacks, 2004). Event segmentation has not previously been considered in relation to the "subjective present". However, if behaviour is actually structured in temporal units that correspond to and reflect the 3 s temporal integration interval, as has been claimed (Pöppel, 2009; Wittmann, 2011), then it follows that structures on that scale should be comprehended in perception of behaviour. Thus, it follows from the hypothesis of a 3 s "subjective present" that event segmentation should have a temporal resolution of about 3 s.

In several studies, participants have been asked to segment behaviour sequences into either the smallest units that seemed natural to them or the largest. Mean durations of the largest constructed segments varied from 14.1 s in the first study of this kind, by Newtson (1973), to 42.0 s in a study by Hard, Recchia, and Tversky (2011), but it is the time scale of the smallest units that is of most relevance to the present hypothesis. Results of studies that asked participants to segment on the smallest natural scale and that reported mean segment durations are summarised in Table 1.

The table shows a range of mean segment durations from 3.5 s (Wilder, 1978a) to 19.1 s (Swallow, Zacks, & Abrams, 2009). The shortest mean segmentations occurred with stimuli depicting solo actors in highly structured activities. Wilder (1978a) used a film of an actor constructing booklets by picking up sheets of paper from different piles one after another, stapling them together, and placing the completed booklet on a table, repeating this sequence 20 times at a pace of 15 s per assembly. The longest mean segmentations occurred with stimuli depicting multiple actors in complex, busy activities in Hollywood movies (Swallow et al., 2009).

Although the shortest mean segmentations have durations close to those of the 3 s "subjective present" (Newtson, 1973; Wilder, 1978a, 1978b), most studies found much longer mean segmentations. The means disguise a wide and skewed distribution: Zacks, Tversky, and Iyer (2001) found segments ranging up to 55 s in length. These results do not support the hypothesis of temporal integration on the time scale of the hypothesized "subjective present". It is important to bear in mind that the standard instructions given to participants were to segment the stimulus materials into the shortest units that seemed meaningful to them; yet, in almost all cases, that shortest meaningful unit was substantially longer than the 3 s "subjective present". Swallow et al. (2009) set recognition memory tests for their participants and found that memory was significantly affected by the occurrence of event boundaries; for example, memory for information that was carried across an event boundary before being tested was poorer than memory for information on the same time scale but within the same event. A detailed discussion of their results is outside the scope of this paper, but the results support the hypothesis of temporal integration of event information that is variable on a scale from a few seconds to a minute: segmentation marks meaningful boundaries between events, which are temporally integrated units of social behaviour, and this aids comprehension and memory (see also Pettijohn, Thompson, Tamplin, Krawietz, & Radvansky, 2016). This has implications for the hypothesized "subjective present": if the "subjective present" is a product of temporal integration, and if the time scale of temporal integration in perception of behaviour varies from a few seconds to a minute, then the "subjective present" must vary on the same time scale.

Reversals of perception in bistable ambiguous figures and binocular rivalry stimuli

Most people are familiar with the Necker cube, a two-dimensional outline of a cube that can be perceived as being in one of two orientations. Usually just one orientation is perceived at any given time, but the percept shifts back and forth from one to the other. This is an example of a bistable ambiguous figure. Pöppel (1997, 2009) and Wittmann (2011) claimed that shifts from one percept to the other occur on a time scale of approximately 3 s, and that this is another case of temporal integration on the 3 s time scale. Perceptual shifts can also be obtained by presenting different images that cannot be fused, one to each eye: the percept tends to shift back and forth between one image and the other. That is binocular rivalry. There have been numerous research studies of perceptual reversals; here I focus on data concerning the time scale of reversals.

Table 2 presents mean reversal intervals for percepts of the Necker cube reported in several experiments. The table shows reasonable consistency across experiments, with a range from 4.4 s to 7.5 s. In addition, Gomez, Argandoña, Solier, Angulo, and Váquez (1995) presented a binocular rivalry stimulus and found a mean reversal interval of 7.31 s. Kogo, Hermans, Stuer, van Ee, & Wagemans (2015) used Rubin's (1921) figure that can be seen as a face or as a vase and found a mean reversal interval of 5.5 s. They also used an anomalous transparency figure created by Kanizsa (1979) and found a mean reversal interval of 7.5 s. These means were significantly different. Radilová and Pöppel (1990) used a figure called the Schröder staircase in which a series of steps can be perceived in normal orientation or inverted, and found a mean percept duration of about 4 s for the normal orientation and about 2.5 s for the inverted orientation. These durations increased to about 5.3 and 3 s, respectively, if the figure was presented as a photographic negative. Gori, Hamburger, and Spillmann (2006) presented a static figure that gives rise to illusory impressions of clockwise or counter-clockwise rotation. They found mean durations of 5.0 s for clockwise and 4.4 s for counter-clockwise rotation percepts. Ilg, Burazanis, Wohlschläger, Wöller, Wagenpfeil, and Mühlau (2008) presented an ambiguous apparent motion stimulus. The stimulus is a circle of 12 dots with successive presentations in which the dots are rotated by an amount equal to their radius. This can be perceived as rotation in either direction. They manipulated frequency of alternation in the stimulus with 5.0, 6.0, or 7.5 frames per s, and percept duration increased from 2.8 s at 7.5 frames per s to 3.4 s at 5.0 frames per s.

The results of the studies by Radilová and Pöppel (1990) and Ilg et al. (2008) are consistent with the hypothesis of a 3 s "subjective present". The other findings all suggest percept durations a little longer than that. It is perhaps pertinent that Table 2 reports means. Gomez et al. (1995) found that the distribution of reversal intervals was skewed, with a modal value lower than the mean, in the range 2 - 4 s. On the one hand, this suggests that the mode gives some support to the 3 s "subjective present" hypothesis. On the other hand, under the 3 s "subjective present" hypothesis, one would expect a distribution where most intervals fall between the hypothesized lower and upper bounds of the "subjective present". In that case, the mode and the mean should both be lower than the upper temporal bound of the "subjective present". Instead, the mean is well above it and the mode approximately at the upper bound. Only a small proportion of reversal intervals occur at any value lower than that, whereas a much higher proportion occur at intervals longer than the 3 s upper bound, which is why the mean is higher than the mode. Arguably, if the distribution of reversal intervals is skewed, one would expect the opposite direction of skew from what is observed, with hardly any intervals longer than the hypothesized upper bound.

Some additional features of the research are pertinent. Orbach, Ehrlich, & Heath (1963) and Orbach, Zucker, and Olson (1966) used rapid intermittent presentations of the Necker cube with varying amounts of time on and off, and found that this had a substantial effect on percept duration. Orbach et al. (1966) found mean percept durations of 4.3 s at 10 ms off, 1.6 s (the shortest duration recorded) at 200 ms off, and 12.5 s at 800 ms off. For on-time manipulations, percept duration was about 20 s with short on-time and about 3 s with long on-time. Meng and Tong (2004) found that observers could influence the duration of percepts by means of attention, both increasing the duration of the attended percept and decreasing that of the unattended percept. Scotto (1991) presented the Necker cube moving back and forth and participants tracked it with pursuit eye movements. Changes in the percept were entrained to the time of direction reversals, which happened at intervals of 1.3 s. Long, Toppino, & Mondin (1992) and Toppino and Long (2015) used pre-exposure to an unambiguous version of the figure (i.e. a version that only fitted one of the two percepts) and found substantial effects of that on percept duration when the ambiguous figure was presented. Long et al. (1992) found the same percept as in the unambiguous stimulus with short pre-exposure periods and the opposite percept with long pre-exposure periods. Toppino and Long (2015) used long pre-exposure presentation and found that mean percept duration was increased for the percept opposite to the preexposure stimulus but not affected for the percept that was the same as the preexposure stimulus.

There are also findings that indicate that the mechanisms underlying perceptual reversal are to some degree specific to modality, or even to the specific kind of stimulus used. Wernery, Atmanspacher, Kornmeier, Candia, Folkers, and Wittmann (2015) compared reversals with the Necker cube to reversals with an auditory stimulus, a repeating sequence of the syllables "au" and "gen" which in German can be heard as "augen" (eyes) or "genau" (exact). They found no significant correlation across participants between reversal rates for the visual and auditory stimuli. This supports a hypothesis that visual and auditory reversals are determined by separate mechanisms. Strüber and Stadler (1999) compared perceptual reversals for figures where the reversal involves realignment of the reference frame or orientation, such as the Necker cube, with figures where the reversal involves transformation in semantic features of the percept, such as the figure that can be perceived as either a duck or a rabbit. They found a consistently greater degree of control over voluntary reversals for the latter kind of figure than for the former. This implies separate mechanisms for different kinds of visual ambiguous figures, with different degrees of involvement of cognitive or attentional factors. The evidence of these two studies is disconfirmatory for the hypothesis that perceptual reversals indicate the time scale of a common temporal integration mechanism for the "subjective present".

In summary, the mean percept duration for the Necker cube is about 5 s. This is beyond the upper limit of the 3 s "subjective present". A majority of results for perceptual reversals with other ambiguous figures and binocular rivalry stimuli also show means longer than 3 s. The hypothesis of temporal integration as a mechanism for determining the "subjective present" does not predict several features of the results, including the alteration of percept duration (in particular increased to as much as 20 s) by intermittent presentation, by manipulations of attention, by object motion, and by pre-exposure stimuli. Most important, the hypothesis does not explain consistent differences in mean reversal times between different ambiguous figures and between modalities. Several findings indicate that adaptation plays a role in shifts of perception: in simple terms, neurons for one interpretation become adapted or fatigued to a point where they are less excited than neurons for the alternative interpretation, which then take over (Gomez et al., 1995). Recent papers have developed theoretical models of reversability in bistable percepts that have proposed processes specific to the phenomenon (Furstenau, 2014; Lee, 2014; and see also Holcombe & Seizova-Cajic, 2008). Thus, the time scale of, and effects of manipulated variables on, perceptual reversals reflects, not a general temporal integration process constructing the "subjective present", but a specific feature of visual processing (fatigue or adaptation) that happens to occur on a similar (though not identical) time scale.

Inhibition of return (IOR)

When a salient, attention-capturing stimulus is briefly presented at a given location in the peripheral visual field, perceptual processing of a subsequent stimulus in the same location is impaired for a period of time. This is known as "inhibition of return" (IOR) (Samuel & Kat, 2003). Pöppel (2009) referred to this phenomenon as support for the 3 s "subjective present" on the grounds that the duration of the inhibitory period is about 3 s.

Samuel and Kat (2003) reviewed the literature on IOR. They showed that IOR does not occur at SOA < 300 ms, and that the opposite of inhibition occurs on that time scale. There was no determined upper limit because no studies had presented SOA > 3000 ms. In their own study they tested SOAs up to 4200 ms, and found that IOR did not occur at durations longer than 3000 ms. This finding was later replicated by Bao, Wang, Liang, Wang, Pöppel, and Li (2013).

On grounds of time scale alone, this evidence renders IOR consistent with the hypothesis of a 3 s time window for the "subjective present". However, Dodd and Pratt (2007) presented a series of stimuli at multiple locations instead of just one, followed by a target that participants were instructed to detect and respond to as quickly as possible. With this multiple stimulus presentation, they found IOR for durations up to 6000 ms. Dodd, van der Stigehel, and Hollingsworth (2009) pointed out that IOR had previously only been demonstrated with visual search tasks, so they used a visual search task, a memorisation task, a pleasantness rating task, and free viewing in different conditions. They found IOR only in the visual search task, and in fact in the other tasks facilitation, the opposite of IOR, occurred. As the authors concluded, this indicates that IOR is a task-specific phenomenon, and for that reason it cannot be interpreted as evidence for a general "subjective present" window of 3 s.

Mismatch negativity (MMN)

If a series of identical auditory or visual stimuli is presented at regular intervals, event-related potential (ERP) recordings show brain responses occurring if a different, unexpected ("deviant") stimulus is presented at some point in the series. These brain responses have become known as mismatch negativity (MMN). Pöppel (2009) discussed a study by Sams, Hari, Rif, and Knuutila (1993) in which it was reported that the strongest MMN occurred with SOA = 3 s. Pöppel argued that "approximately every 3 s the sensory channel is more sensitive than at other times for new information coming from the external or internal environment" (p. 1893). This was taken as supporting the hypothesis of a temporal window of integration on that time scale.

There have been hundreds of investigations of the MMN in both the auditory and visual modalities (see, for example, Bartha-Doering, Deuster, Giordano, Zehnhoff-Dinnesen, & Dobel, 2015; Näätänen, Astikainen, Ruusuvirta, & Huotilainen, 2010; Näätänen, Paavilainen, Rinne, & Alho, 2007; Stefanics, Kremláček, & Czigler, 2014). Only a few studies have investigated the upper limit of SOA that can elicit the MMN. The study by Sams et al. (1993) is one of those. The very title of their paper should be a clue that the results did not support the hypothesis of the 3 s "subjective present": "The human auditory sensory memory trace persists about 10 sec" (p. 363). Although they did report the strongest MMN at SOA = 3 s, they also reported significant MMN occurring at SOA up to 9 s, though not at 12 s, the longest value used. However, Ritter, Sussman, Molholm, and Foxe (2002) pointed out that four of the seven participants in the study by Sams et al. did show a significant MMN at SOA = 12 s. Moreover, Sams et al. (1993) themselves pointed out that the reason for the peak response at SOA = 3 s is just that MMN amplitude depends on the interval between the <u>deviant</u> stimuli, which is determined by but much longer than the SOA. Bottcher-Gandor and Ullspacher (1992) reported MMN occurring with SOA = 1, 6, and 10 s, and with similar amplitudes for all three values. Wang, Lin, Zhou, Pöppel, and Bao (2015) used auditory stimuli with SOA ranging from 1.5 to 6 s. They analysed data from four electrodes. One showed significant difference in response between SOA = 1.5 and 3 s and SOA = 4.5 and 6 s. The other electrodes showed either a general decline in MMN amplitude as SOA increased or no obvious pattern. Overall, findings that the MMN occurs with SOA in the range 9 - 12 s rules out the hypothesis that it is a manifestation of temporal integration on a time scale around 3 s. It should also be noted that the MMN is a response to a stimulus that differs from a preceding series of similar stimuli. To detect a difference requires learning of the series. If we assume that at least three repetitions are necessary to establish an expectation, SOA = 3 s actually implies temporal integration (learning) on a much longer time scale. So it is not clear that findings of MMN research can even speak to the issue of the 3 s "subjective present".

Overview of the literature

On close examination, the research reviewed here provides little if any evidence for a general 3 s window of temporal integration that constitutes a "subjective present". Most of the phenomena studied show ranges that extend well beyond 3 s and do not indicate any special significance for the 3 s duration; the research in the previous section on the MMN, showing the MMN occurring at SOA up to 12 s, is an apt example. Previous publications have focussed on means and have neglected variance, which is often substantial: a mean of 3 s does not support the hypothesis of a 3 s "subjective present" if the distribution of values extends up to 10 s or more, as has been found in studies of time units in speech and other behaviour, event segmentation, and the MMN.

Some of this variation could reflect the influence of attentive processing on the phenomena being investigated. This is most obviously the case in sensorimotor synchronisation, where synchronisation performance is affected by voluntary changes of strategy which presumably change the involvement of attention. However, this possibility is not favourable to the "subjective present" hypothesis because the temporal integration involved in the "subjective present" is supposed to be automatic, and would not, therefore, be affected by attentive involvement. If phenomena are subject to influence from attention, or affected by cognitive load, then those phenomena do not represent automatic temporal integration and are not evidence pertaining to the hypothesis of the "subjective present".

Two phenomena do show consistent durations, but they do not match the 3 s duration. There is a limit on subjective rhythmization of about 1500 ms, which is shorter than the 3 s window. The mean percept duration for bistable figures such as the Necker cube is about 5 s, which is longer than the 3 s proposed window and more than three times longer than the limit on subjective rhythmization. Consistency in the time scale of temporal integration across different kinds of processing is a key requirement for the hypothesis of a general temporal window for the "subjective present", so the evidence of lack of consistency is disconfirmatory for that hypothesis.

Although the "subjective present" is hypothesized to be a window of temporal integration of information, little of the research that has been claimed as support for it actually concerns temporal integration. Duration reproduction and sensorimotor synchronization are concerned with timing mechanisms and with memory, not with temporal integration. Behaviour units are concerned with motor output, not perceptual input: that is why I argued that event segmentation, which has to do with perception of behaviour, would be more relevant to temporal integration in the "subjective present". Perceptual reversals are probably the product of low-level visual phenomena such as adaptation and do not reflect any kind of temporal integration. Inhibition of return specifically concerns visual search and does not involve temporal integration. Mismatch negativity is a response to an unpredicted deviation from a series of predictable events. It does involve memory, but on a much longer time scale than 3 s. Thus, although research on the topics discussed here has been treated as relevant to the "subjective present" by previous authors (Montemayor & Wittmann, 2014;

Pöppel, 1997, 2009; Wittmann, 1999, 2009, 2011), this does not appear to be the case for most of them.

Very few studies have been directly concerned with temporal integration on a time scale of seconds. There are two studies that show temporal integration of motion information on a time scale up to 2.8 - 3 s. Burr and Santoro (2001) presented moving random dot patterns embedded in noise. They found that motion coherence sensitivity increased with duration up to 2 - 3 s, consistent with the hypothesis of temporal integration of global motion signals on a time scale up to 3 s. Neri, Morrone, and Burr (1998) presented point-light walker stimuli in the presence of visual noise and found evidence for temporal integration on a time scale up to 2.8 s. Both studies are essentially concerned with detection of signal in noise, so it is likely that the temporal integration is a form of summation that generates a percept when it reaches some criterion level of signal detection (e.g. Ossmy, Moran, Pfeffer, Tsetsos, Usher, & Donner, 2013; Tsetsos, Pfeffer, Jentgens, & Donner, 2015). No other perceptual processing is known to involve integration over such a long time period, so it is unlikely that it reveals a general temporal integration mechanism on that time scale.

Fairhall, Albi, and Melcher (2014) argued that, "[a]ccording to the logic of a temporal window, the brain should be able to combine information within the limits of a single TIW [temporal integration window], even when the order of that information is scrambled, but have difficulty when information is scrambled over longer timescales" (p. 2). They scrambled film clips (video only) within windows with durations varying from 800 to 12800 ms. Participants rated how difficult the clips were to follow. There was a sharp increase in rated difficulty between 1600 and 3200 ms, compared to more gradual changes between all other adjacent window sizes. They ran a second study with windows of 1200, 2000, 2800, 3600, and 4400 ms. This time there was a sharp change between 2000 and 3200 ms with more gradual or no change elsewhere. These results are consistent with a temporal integration window somewhere between 2000 and 3200 ms in duration. The results of this study are at least suggestive, and are the right kind of research to be doing, but there is a need for further investigation in other modalities and with other kinds of stimuli before a case for a general temporal window of around 3 s can be made with confidence.

Temporal integration occurs on many time scales in different processes. There are, as I shall show in the next section, multiple time scales of integration in auditory processing of speech (Lerner, Honey, Silbert, & Hasson, 2011; Stephens et al., 2013). Given that, it would be surprising if no kind of temporal integration occurred on a time scale somewhere near 2 - 3 s. But for a 2 - 3 s window of temporal integration to be a general property of perceptual processing that suffices to define the "subjective present", there would at least have to be a distinctive cluster of different kinds of temporal integration occurring at that time scale, as compared to other time scales. This review shows that there is, at present, no evidence for that.

Reconstructing the "subjective present": the envelope of integration

The review so far has shown that the research evidence does not support the notion of a "subjective present" defined in terms of temporal integration of information on a time scale of a few seconds. The different processes that have been called on as support for the proposal have different characteristic time scales that do not match each other or, mostly, the hypothesized time scale for the "subjective present". Perhaps more important, the processes are, mostly, not temporal integration processes.

Does this mean that there is no "subjective present"? Common experience seems to suggest that there is more to our understanding of what is going on than just the products of perceptual processes. When I play badminton, I am focussed on the present motion of the shuttlecock, on my own action plans, and on what my opponent is doing. But other information is active as well: I know who I am, strengths and weaknesses of my opponents, the rules of the game, the current score, and many other relevant things. These form a nimbus of activated information around the sharp focus of the present moment. But it is not clear that they involve temporal integration. They are just a collection of relevant pieces of information that, for example, may influence my choice of what shot to play next. They resemble more the notion of "mental presence" proposed by Wittmann (2011) as involving "integrating mental processes and enabling conscious experience of a narrative self that has personal identity and continuity over time" (p. 5). Wittmann (2011, p. 5) continued: "Essentially, it is working memory, a system of limited attentional capacity, supplemented by visuospatial, episodic, and phonologic storage systems, which holds information for temporal storage and manipulation". Is there anything more than that, or is that all there is to the "subjective present"?

It remains to be shown that there is a sharp, definable boundary between temporal integration processes and the more general sense of mental presence described by Wittmann (2011). However, there is evidence that temporal integration processes occur beyond what is usually considered the boundaries of perceptual processing, and it can be argued that the products of those processes form at least part of our moment-by-moment understanding of what is going on. In this section I offer some suggestions about what might be involved. Specifically, I propose that, to the extent that there is a "subjective present" at all, it results from the operation of local temporal integration processes each with their own operating characteristics, that the temporal extent of the "subjective present" is set by the density of information and the operating characteristics of individual processes, not by time, and that there is some form of global temporal integration of the products of local integration processes.

I noted earlier that Pöppel (1997, 2009) used the term "subjective present", while Wittmann (2011) preferred the term "experienced moment". Each of these implies a very narrow temporal window: the present is, strictly speaking, just the infinitesimally brief moment separating the past from the future. There may be no entirely satisfactory term, but I shall propose the term "envelope of integration": "envelope" refers to the global coherence of information and "integration" refers to the contents of the envelope as the products of integration processes (both spatial and temporal, though temporal integration is the main focus of this paper).

Local temporal integration and information density

Temporal integration involves different pieces of information or content being related in some way. Two pieces of information are not bound or integrated by the mere fact of being in a memory store at the same time. Let me give two examples of temporal integration on the millisecond time scale. One is integration of sound input on a time scale of 150 - 300 ms to yield a percept of a phoneme (Poeppel, 2003). The other is cross-modal synchronisation to generate a coherent audiovisual percept of a person talking. Both of these operate on a time scale of less than 300 ms (Chen & Vroomen, 2013; Poeppel, 2003). In each case the product of the temporal integration process is a unified percept. It is easy to see why temporal integration might operate on that time scale, because it is part of the general activity of constructing a coherent multimodal perceptual world. Each of the examples concerns local temporal integration, dealing with only a small proportion of the total perceptual information being processed at any one time. There is also a need for more global integration, and the evident overall coherence of the perceptual world shows that it does happen, though little is known about how (Zimmermann, Morrone, & Burr, 2014).

A striking feature of the research reviewed in the present paper is that each process or mechanism has its own operating characteristics and its own time scale. Take the study by Burr and Santoro (2001). This research showed temporal integration of motion information over 2 - 3 s under conditions of noise. This form of temporal integration has a specific task, which is the detection of global (coherent) motion. The time it takes is roughly the time taken to distinguish signal from noise, up to an upper limit that appears to be about 3 s. So this form of temporal integration has what might be called a local processing objective, and the time scale of integration varies in a way that is functionally related to that processing objective. That is likely to be a general truth about temporal integration: it is in the service of some specific processing objective, and its time scale is influenced by (i) any upper temporal limit there might be on its integration capacity; (ii) the time taken to accomplish the processing objective, which itself is determined by such factors as signal-to-noise ratio; (iii) operating characteristics of the particular process or mechanism in question; (iv) where information is drawn from working memory, working memory capacity limits, including demands on cognitive capacity from other concurrent processes.

Temporal integration occurs on multiple hierarchical levels. Lacking space for a comprehensive review, I shall focus on speech perception. On short time scales, two temporal windows of integration in speech perception have been identified, approximately 20 - 50 ms and 150 - 300 ms (Chait, Greenberg, Arai, Simon, & Poeppel, 2015; Poeppel, 2003). On longer time scales, several studies have investigated temporal integration of speech using the method of scrambling on selected time scales (Hasson, Chen, & Honey, 2015; Honey, Thesen, Donner, Silbert, Carlson, Devinsky, Doyle, Rubin, Heeger, & Hasson, 2012; Lerner et al., 2011; Lerner, Honey, Katkov, & Hasson, 2014; Stephens et al., 2013). The time scales of scrambling have been quite coarsely differentiated; for example Stephens et al. (2013) scrambled an audio recording of a story reading on three time scales of 0.7 ± 0.5 s (words), 7.7 ± 3.5 s (sentences), and 38.1 ± 17.6 s (paragraphs). However, two important findings have emerged from these studies. One is support for multiple temporal integration windows for speech, shown by responses to different time scales of scrambling in different areas of the brain. Hasson et al. (2015), summarising the previous studies, proposed five time scales ranging from phonemes at the millisecond time scale through words, sentences, paragraphs, and narrative.

The other important finding is that temporal windows of integration are scaled to informational units and not temporal ones. Lerner et al. (2014) tested this by manipulating speed of presentation of an auditory verbal narrative. They found evidence that temporal integration was rescaled both at short time scales in early auditory processing and at longer time scales corresponding to larger linguistic units. They found also that temporal rescaling "started to break down for stimuli presented at double speed, and intelligibility was also impaired for these stimuli" (p. 2014). This is evidence that intelligibility depends on temporal integration of informational units, such that speech ceases to be intelligible if the rate of presentation of information exceeds the capacity of a given level of temporal integration.

Two conclusions can be drawn from this research. One is that there is no single time window defining the envelope of integration. Instead, there are multiple time windows of integration in perceptual processing and beyond, starting on the millisecond time scale and stretching out to the level of narrative on a time scale perhaps as much as hundreds of seconds (Stephens et al., 2013; Hasson et al., 2015). The other is that each window of temporal integration should be conceptualised in terms of information capacity, not time scale. There are limits on the flexibility of the time scale for each window (Lerner et al., 2014), but it is not appropriate to define windows of integration in terms of specific times. It should be noted that the findings I have discussed are specific to auditory linguistic stimuli, and it is likely that other kinds of processing have different temporal integration windows suitable to the information units with which they deal (e.g. Hasson, Yang, Vaillines, Heeger, & Rubin, 2008). In music perception, Farbood, Heeger, Marcus, Hasson, and Lerner (2015) used a similar scrambling method to present segments on music on three different time scales with mean durations of 1.29 s, 6.32 s, and 38.28 s. They found different patterns of brain activation for the different time scales, indicating different mechanisms for integrating musical information on different time scales. These are similar to the time scales identified by Stephens et al. (2013), so it is possible that similar temporal windows of integration obtain in both speech and music perception. However, greater variation in the time scales used for scrambling the stimulus materials would be required before this conclusion can be accepted.

In addition, integration is not exclusively on the temporal dimension, whether it is defined in terms of time or information density. There is also spatial integration, which includes bounded object individuation and identification (Kahneman, Treisman, & Gibbs, 1992; Wutz & Melcher, 2014), scene gist analysis (Greene & Oliva, 2009; Groen, Ghebreab, Prins, Lamme, & Scholte, 2013), and a spatiotopic map that integrates information across the whole perceived environment (Yoshimoto, Uchida-Ota, & Takeuchi, 2014; Zimmermann et al., 2014). There is also what may be called semantic integration which becomes increasingly important as information proceeds down the processing stream into working memory (Sachs, 1967, 1974). Semantic integration may encompass multiple time scales, as in the example of speech perception, with all levels being integrated into a composite semantic representation of what is going on (e.g. what the speaker is talking about).

Thus, the envelope of integration is a spatially, temporally, and semantically organised and integrated representation of what is going on, combining local integration of all three kinds into a coherent overall representation. It is not defined in terms of time scales. Local processes have their own operating characteristics which may impose limits on the temporal extent of information that can be incorporated in the generation of a single informational product, but there is no global temporal extent that characterises all of the envelope of integration. In this respect it is inappropriate to define a subjective present or experienced moment in terms of time, even a range of times.

Loss and qualitative transformation of information

Early visual processing has very large capacity. Once perceptual construction is complete or at least fairly advanced, retention of information about recent percepts is subject to increasing capacity limits. Jacob, Breitmeyer, and Treviño (2013) found evidence for three stages of visual processing. The times given here are SOAs, where the first stimulus is a prime and the second is a probe, and participants have to judge whether the probe is the same as the prime or not. The first processing stage ran from 0 ms to 133 ms SOA, and the authors identified it with visible persistence. This represents very short-lived retention of otherwise transient information, possibly to facilitate further processing of it (Coltheart, 1980; di Lollo, 1977; Wutz & Melcher, 2014).

The second stage ran from 240 ms to 720 ms, and the authors identified this with iconic visible persistence. This refers to a brief but high-capacity visual store first researched by Sperling (1960). If 240 ms marks the onset of storage at that level, that would coincide with the time scale of synchronisation processes, and would suggest that iconic visible persistence stores essentially fully temporally integrated visual information. Information in iconic store decays rapidly to the limited capacity of visual short-term memory. Much of the decay occurs in the first 250 ms, but decay may not be complete until about 1000 ms (Demkiw & Michaels, 1976; Schill & Zetsche, 1995; Sperling, 1960).

The third stage ran from 1200 ms onwards, and Jacob et al. identified it with visual short-term memory. Their results showed limited capacity with rapid decay. It may not be possible to specify a decay function for visual short-term memory, partly because of the involvement of attention and interference effects (e.g. Luck & Vogel, 2013; Rensink, 2002; Oberauer & Lin, 2017), but in the absence of active attentive maintenance of information, loss of information occurs quickly, as many studies of change blindness have shown (Rensink, 2002)

The hypothesis of these three successive information stores was further supported by a study by Öğmen, Ekiz, Huynh, Bedell, and Tripathy (2013). Although their research did not pin down the timing of the storage systems, their properties appear to correspond to those of the three stores identified by Jacob et al. (2013). Thus, the first stage may correspond to visible persistence, but in any case is characterised by high capacity. There is then a bottleneck to the second stage, which corresponds to iconic informational persistence, and the results of the study by Öğmen et al. suggest that this has substantially lower capacity than the first stage. There is then a further information bottleneck to the third stage, which is visual short-term memory. This in turn has an information bottleneck corresponding to the initial limited capacity of visual short-term memory, and further loss of information thereafter, through decay or interference.

This structure and time course of stores and bottlenecks may be at least partly specific to vision. For example, there have been studies on the equivalent of iconic visible persistence in both the auditory and somatosensory modalities (Auvray, Gallace, & Spence, 2011; Bliss, Crane, Mansfield, & Townsend, 1966; Darwin, Turvey, & Crowder, 1972), and the evidence indicates that these stores have different properties. The auditory equivalent to the iconic store seems to decay over a longer period, about 2000 ms (Darwin et al., 1972), and the somatosensory equivalent seems not to have such a great capacity (Auvray et al., 2011; Bliss et al., 1966).

It should be noted that the various bottlenecks are not necessarily just quantitative. Öğmen et al. (2013) argued that the first bottleneck they identified was associated with precision, which refers to detail in specification of items: they suggested that it is possible to encode several items with low precision but few with high precision. The bottleneck in short-term memory may be primarily a semantic bottleneck. As we have seen, semantic information is retained better than syntactic information over time scales longer than 3 s (Roll et al., 2012, 2013; Sachs, 1967, 1974). That is for speech, however, and it is not clear whether the same would apply to other auditory stimuli or to stimuli in other modalities. Thus, as information progresses through the processing stream, there is both loss of information and change in the predominant mode of representation of information. At risk of oversimplifying the latter, in early stages the representation resembles a visual after-image, an essentially complete representation of surface features of the stimulus (Coltheart, 1980) but, as processing continues, semantic information is gradually added to the representation (e.g. Rossion & Caharel, 2011) while surface information is gradually lost (e.g. Sachs, 1967; Sperling, 1960).

It is also important to note that the stream of processing in perception is not a bottom-up, feedforward stream: abundant research has shown the involvement of pre-existing knowledge in perceptual processing from early stages on, and perception can be regarded as a kind of reiterative, multi-level hypothesis-testing activity where possible interpretations are tested against incoming evidence (e.g. Bar, Kassam, Ghuman, Boshyan, Schmid, Dale, Hämälainen, Marinkovic, Schacter, Rosen, & Halgren, 2006; Clark, 2013; Di Lollo, 2012; Enns & Di Lollo, 2000; Hochstein & Ahissar, 2002; Hohwy, 2013; Kahan & Enns, 2014; Kwon, Tadin, & Knill, 2015; Tapia & Beck, 2014). Transformation of information as processing proceeds should be understood as a manifestation of that hypothesis-testing activity.

I mentioned that the third stage of processing identified by Jacob et al. (2013) corresponds to visual working memory. The time scale that has conventionally been taken as that of the "subjective present", a few seconds, is also that of working memory. Can it be argued, then, that there is a general limit on the time scale of temporal integration in relation to the hypothesized "subjective present" and that it is set by the limited capacity of working memory? This is probably not the case.

One problem is that working memory is associated more with Wittmann's (2011) notion of "mental presence", the nimbus of activated information, such as the narrative self, that surrounds current processing. Wittmann (2011) argued that mental presence covers a series of experienced moments (his term for the "subjective present") and integrates them into a unified experience which, he argued, was based on working memory function. However, he also argued that it has no absolute temporal boundary, which would take it beyond the time scale of working memory.

What that illustrates is that, once information enters working memory, there is no absolute time limit on its retention there. An item of information could, in principle, be retained indefinitely through attentive maintenance and rehearsal (Rensink, 2000). Or it could be transferred to long-term memory and then retrieved at any time thereafter. In that way, some form of temporal integration can occur between an item that has just entered working memory and one that has been there for, say 20 s, or one that was stored in long-term memory an hour ago and has just been retrieved. Temporal integration of that sort may be possible, and indeed it appears consistent with the mental presence notion, but it goes well beyond the sense of a "subjective present".

Moreover, not all processing on a time scale of seconds is bound by the capacity of working memory. Much processing is automatic, and automatic processing is known to have much higher capacity than working memory (Moors & De Houwer, 2006; Shiffrin & Schneider, 1977). The distinction between automatic and non-automatic (often called controlled) processing is not absolutely sharp and clear. It has been proposed, for example, that there is a continuum of automaticity, determined primarily by amount of training on the process or task in question (Logan, 1985).This implies that there is no general capacity limit that determines the upper limit of the time scale of temporal integration. Any given process may comprise component sub-processes that have varying degrees of automaticity, thereby transcending the limited capacity of working memory.

Bearing all that in mind, it is probable that there is no sharp division between the envelope of happening and mental presence. There is likely to be a gradual transition from fully automatic to more controlled processing, as information proceeds through successive stages of processing from initial low level perceptual processing through higher perceptual processing, iconic time scale processing, temporal integration beyond the iconic time scale, and processing of information in working memory. The envelope of happening is generated mainly by temporal integration on the supra-second scale and by processes that are mainly automatic; mental presence is generated by activation of stored information, particularly of a semantic nature, that is integrated into an informational construct of context, and that is mainly associated with controlled processes operating on information in working memory. Attentive processing, automatic processing, and controlled processing are probably all involved in both the envelope of happening and mental presence, but to different degrees and in different ways.

Local temporal integration probably occurs at all stages of the stream of processing, depending on the operating characteristics of individual processes and on information capacity limits at different stages. All temporal integration processes contribute to the content of the envelope of integration, and in all cases the reach of temporal integration is determined by information density combined with the degree of automaticity in the processing. Primarily, therefore, the envelope of integration needs to be understood in terms of the operating characteristics and information capacity of the processes that generate its content.

Global integration

The most evident feature of the envelope of integration is the subjective experience of overall, global coherence. In the case of speech perception, syllables, words, sentences, and narrative are each the products of local informational integration involving different amounts (or time slices) of information. These locally integrated products are also globally integrated to form a coherent and unified percept of what the speaker is saying. That in turn is integrated with the context and whatever else the listener is attending to at the time. The extent of global coherence should not be overestimated. Information that is not attended tends to decay rapidly, even if it is in working memory (Rensink, 2000). Thus, information currently being attended has the subjective experience of global coherence, and this property is assumed to hold for all other information about what is going on, but this is not so much the case for information that is not at the focus of attention. Nevertheless, it is this experience of global coherence that perhaps does more than any other single feature to give the envelope of integration its subjective character.

Global integration is underpinned by local integration. Binding and synchronisation properties operating on the scale up to about 250 ms are central to the construction of a coherent perceptual world. However, given that the processes that I have been discussing in this section, such as the more global end of speech processing, take as their input the units generated by earlier, more local, information-dense processes, it is clear that there must be additional kinds of integration that deal with the outputs of multiple processes operating on the supra-second scale. Although, as in the case of speech perception, a good deal is known about local integration (where, for present purposes, the whole of speech processing is local in the sense of being distinct from, say, visual layout processing or the kinaesthetic body map), there has been little if any research on the integration of different local processing outputs into a globally integrated representation, the envelope of integration as a whole. To the extent that attention has a narrow and local focus, global integration may not be complete. Without it, however, the output of different processes on the supra-second scale would fall apart into separate representations.

The envelope of integration, then, is an integrated representation of the current percept (the experienced present) with information about the recent past, which is constructed on several time scales or levels of information processing, each of which involves its own internal temporal integration, and each of which is temporally integrated into a coherent representation of what is going on. The whole integrated envelope of integration encompasses different modes of representation, but all are integrated into the overall representation.

Outstanding issues

What is proposed here is not a fully developed theoretical proposal, which would be outside the remit of this paper. It is a set of suggestions to guide further research on the issue. Instead of looking for temporal limits on integration of information on the supra-second scale, research could usefully aim to (i) ascertain more of the particular features of local temporal integration processes and their products; (ii) find out more about capacity limits on temporal integration processes, and in particular the role of information density; (iii) investigate not just loss but also selectivity and transformation in information as it proceeds through the processing stream; (iv) investigate the possibility of some form of global integration that binds the products of multifarious processes into a coherent representation of happening on the supra-second time scale. The foregoing analysis suggests that that representation would function as a semantic context for ongoing perceptual processing. There are several important outstanding issues that any comprehensive theory of the envelope of integration would have to address, and which I shall briefly list.

1. If there is temporal integration on a time scale of a few seconds, whether it be the envelope of integration sketched out here or the "subjective present" in the publications of Pöppel and Wittmann and colleagues, it straddles the boundary (to the extent that there is one) between perception and memory, or between perceptual processing and post-perceptual cognitive processing. Bearing in mind also the earlier remarks that perception itself can be understood as involving feedback or reentrant processing and hypothesis testing, the way in which sensory/perceptual input, working memory, pre-existing knowledge, and cognitive

processes such as judgment and decision making interact in temporal integration and in constructing the informational content is a complex issue that demands much closer scrutiny.

2. This review has deliberately avoided the issue of consciousness. Given that there have been proposals about possible functions of consciousness, particularly in relation to working memory (Baars, 1988, 2002; Dehaene & Changeux, 2011; Dehaene & Naccache, 2001), it may be that this issue has to be confronted to make full sense of an envelope of integration or "subjective present".

3. To the extent that there is a coherent body of information on a time scale of a few seconds, there are issues to do with how that body of information is updated. There have been proposals that updating on the millisecond time scale occurs periodically, so that there are discrete, successive temporal frames of experience (Pöppel, 1997, 2009; Geissler & Kompass, 2001; Lehmann, Strik, Henggeler, Koenig, & Koukkou, 1998; VanRullen, Zoefel, & Ilhan, 2014), and there are good reasons to think that synchronisation occurs over large areas of brain function (e.g. Doesburg, Roggeveen, Kitajo, & Ward, 2008; Gregoriou, Paneri, & Sapountzis, 2015; Hanslmayr, Gross, Klimesch, & Shapiro, 2011; Jensen, Gips, Bergmann, & Bonnefond, 2014; Lisman & Jensen, 2013; Roberts, Hsieh, & Ranganath, 2013). Can it be argued, then, that the envelope of happening or the "subjective present" is updated periodically in the manner of successive discrete frames?

With specific reference to the "subjective present", Pöppel (2009) claimed that it segments the stream of consciousness into discrete windows of approximately 3 s. Taken literally, this would suggest that all information is integrated over a time scale of 3 s, which means presumably that it would refer to the most recent 3 s of processing, and then after 3 s it is all completely replaced by the integrated products of the next 3 s of processing. This is implausible. Consider two simple problems for this proposal. One is that two perceptual products separated by just a few milliseconds would not be integrated if they fell on different sides of a boundary between one 3 s unit and the next. Such separation would compromise local processing goals and create informational disjunctions that are likely to have detrimental impact on functioning. The other problem is that total replacement of one 3 s informational window by the next one would imply complete loss of information about the contents of the previous window. This can be dismissed on empirical grounds: as we have seen, temporal integration in speech perception operates on multiple time scales extending far beyond 3 s, and the sense of what is going on in an utterance depends on an integrated representation of information on all of those time scales.

If the notion of a window has any applicability on the proposed time scale of the "subjective present", the "travelling window" hypothesis proposed by Allport (1968) would make more sense. Consider an analogy with a train window. A segment of the outside world is visible through the window. As the train moves, so the view, the segment, is gradually updated by new components entering from one side and old components exiting at the other side. The result is globally coherent view over a specific span, with gradual updating of the view as the train moves on. It is still problematic, however, because it implies complete and essentially arbitrary loss of all information when it reaches the outgoing side of the window. This is again disconfirmed by the speech perception research, and again functional considerations are surely paramount: information is retained and integrated to the extent that it remains useful, not to the extent that it fits an arbitrary temporal window.

Pöppel (2009) cited research by Steriade and colleagues showing oscillations with a frequency of 2 - 3 s and suggested that these might support his proposed pre-semantic temporal integration. However, the oscillations in question only occur during sleep or under anaesthesia (Steriade, Nuñez, & Amica, 1993), so are not candidates for the organisation of information in the waking state. In conclusion to this point, then, it is most likely that there are no discrete frames of experience on the supra-second time scale: that information is integrated and retained so long as it is useful, and that information is updated as and when new products of local temporal integration processes happen to emerge, which can be at any time.

4. Although there is abundant evidence that temporal integration occurs, at least locally, mechanisms of temporal integration remain obscure, especially in connection with the hypothesized global integration. Any integration of information across time requires retention of information about the recent past, but integration is much more than just the retention of information about the past few seconds: it involves, for example, generation of a coherent summary impression of the course of events. In Burr and Santoro (2001) there is evidence for temporal integration as a means of identifying a motion signal in noisy conditions; in the speech perception research there is evidence for construction of a word by integration of information across its time course (Poeppel, 2003). The specific mechanisms involved must differ because the tasks they execute differ, but little is known about how those and other temporal integration mechanisms work.

Conclusion

The idea of a three-second "subjective present" is simple and eye-catching, which probably contributes to its popularity. The truth is far from simple. There is no three-second subjective present or experienced moment. There is, instead, an envelope of integration in which there are multiple qualitatively different representations of what is going on, each occupying different and variable time scales of integration, and each integrated into an overall representation that has no fixed time limit, but in any case extends well beyond the three second mark. Specific content of the envelope of integration is updated as and when the relevant processes produce new information: because the specific processes operate on different time scales (and are governed by informational capacity limits rather than time limits), it cannot be the case that all contents of the envelope of integration are updated at one time, in lockstep. Instead, piecemeal updating is accompanied by integration processes that ensure that our perception of what is going on remains coherent.

Footnote

1. Varela (1999) postulated three scales of duration, one of which corresponds approximately to the three-second subjective present. This is the "relaxation time for large-scale integration (the '1' scale)" (p. 117). Varela argued that this scale relates to the function of cell assemblies and defined a cell assembly as "a distributed subset of neurons with strong reciprocal connections" (p. 117). He stated that a cell assembly "must have a *relaxation time* followed by a bifurcation or phase transition, that is, a time of emergence within which it arises, flourishes, and subsides, only to begin another cycle" (p. 117). He wrote that this "holding time... must be comparable to the time it takes for a cognitive act to be completed, i.e. on the order of a few seconds, the 1 scale" (p. 117). He referred to "synchronous coupling of neuronal assemblies; neuronal-level constitutive events, which have a duration on the 1/10 scale, forming aggregates that manifest as incompressible but complete cognitive acts on the 1 scale" (p. 118). Despite the name, this time scale is not precisely 1 s. Varela mentioned evidence of duration estimation being precise up to 2 - 3 s but not beyond it and spontaneous speech and intentional movements being organised in units on a scale of 2 - 3 s, so it would appear that 2 - 3 s is compatible with the '1' scale. However, no published studies were cited in support of those claims.

References

Akyürek, E. G., & Wolff, M. J. (2016). Extended temporal integration in rapid serial visual presentation: attentional control at Lag 1 and beyond. *Acta Psychologica, 168,* 50-64.

Allport, D. A. (1968). Phenomenal simultaneity and the perceptual moment hypothesis. *British Journal of Psychology, 59,* 395-406.

Auvray, M., Gallace, A., & Spence, C. (2011). Tactile short-term memory for stimuli presented on the fingertips and across the rest of the body surface. *Attention, Perception, and Psychophysics, 73,* 1227-1241.

Baars, B. J. (1988). *A Cognitive Theory of Consciousness.* Cambridge University Press.

Baars, B. J. (2002). The conscious access hypothesis: origins and recent evidence. *Trends in Cognitive Sciences, 6,* 47-52.

Baath, R. (2015). Subjective rhythmization: a replication and an assessment of two theoretical explanations. *Music Perception, 33,* 244-254.

Bao, Y., Wang, Z., Liang, W., Wang, Y., Pöppel, E., & Li, H. (2013). Inhibition of return at different eccentricities in the visual field share the same temporal window. *Neuroscience Letters, 534,* 7-11.

Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., Hämälainen, M. S., Marinkovic, K., Schacter, D. L., Rosen, B. R., & Halgren, E. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of America, 103,* 449-454.

Bartha-Doering, L., Deuster, D., Giordano, V., Zehnhoff-Dinnesen, A., & Dobel, C. (2015). A systematic review of the mismatch negativity as an index for auditory sensory memory: from basic research to clinical and developmental perspectives. *Psychophysiology, 52,* 1115-1130.

Blakemore, S.-J., Frith, C. D., & Wolpert, D. W. (1999). Spatiotemporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience, 11,* 551-559.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences, 6,* 237-242.

Bliss, J. C., Crane, H. D., Mansfield, P. K., & Townsend, J. T. (1966). *Perception and Psychophysics, 1,* 273-283.

Block, R. A., & Grondin, S. (2014). Timing and time perception: a selective review and commentary on recent reviews. *Frontiers in Psychology, 5, No. 648.*

Bolton, T. L. (1894). Rhythm. *American Journal of Psychology, 6,* 145-238.

Brown, S. W. (1997). Attentional resources in timing: interference effects in concurrent temporal and nontemporal working memory tasks. *Perception and Psychophysics, 59,* 1118-1140.

Burr, D. C., & Santoro, L. (2001). Temporal integration of optic clow, measured by contrast and coherence thresholds. *Vision Research, 41,* 1891-1899.

Chait, M., Greenberg, S., Arai, T., Simon, J. Z., & Poeppel, D. (2015). Multi-time resolution analysis of speech: evidence from psychophysics. *Frontiers in Neuroscience, 9,* No. 214.

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Attention, Perception, & Psychophysics, 75,* 790-811.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36,* 181+.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27,* 183-228.

Cos, I., Girard, B., & Guignon, E. (2015). Balancing out dwelling and moving: optimal sensorimotor synchronization. *Journal of Neurophysiology, 114,* 146-158.

Darwin, C. J., Turvey, M. T., & Crowder, R. G. (1972). An auditory analogue of the Sperling partial report procedure: evidence for brief auditory storage. *Cognitive Psychology, 3,* 255-267.

Dehaene, S., & Changeux, J.-P. (2011). Experimental and theoretical approaches to conscious processing. *Neuron, 70,* 200-227.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition, 79,* 1-37.

Demkiw, P., & Michaels, C. F. (1976). Motion information in iconic memory. *Acta Psychologica, 40,* 257-264.

Di Lollo, V. (2012). The feature-binding problem is an ill-posed problem. *Trends in Cognitive Sciences, 16,* 317-321.

Dodd, M. D., & Pratt, J. (2007). Rapid onset and long-term inhibition of return in the multiple cuing paradigm. *Psychological Research, 71,* 576-582.

Dodd, M. D., van der Stigchel, S., & Hollingworth, A. (2009). Inhibition of return and facilitation of return as a function of visual task. *Psychological Science, 20,* 333-339.

Doesburg, S. M., Roggeveen, A. B., Kitajo, K., & Ward, L. M. (2008). Large-scale gamma-band phase synchronization and selective attention. *Cerebral Cortex, 18,* 386-396.

Elbert, T., Ulrich, R., Rockstroh, B., & Lutzenberger, W. (1991). The processing of temporal intervals reflected by CNV-like brain potentials. *Psychophysiology, 28,* 648-655.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345-352.

Fairhall, S. L., Albi, A., & Melcher, D. (2014). Temporal integration windows for naturalistic visual sequences. *Plos One, 9,* e102248.

Farbood, M. M., Heeger, D. J., Marcus, G., Hasson, U., & Lerner, Y. (2015). The neural processing of hierarchical structure in music and speech at different timescales. *Frontiers in Neuroscience, 9,* No. 157.

Feldhütter, I., Schleidt, M., & Eibl-Eibesfeldt, I. (1990). Moving in the beat of seconds: analysis of the time structure of human actions. *Ethology and Sociobiology, 11,* 511-520.

Flanagan, J. R., & Wing, A. M. (1997). The role of internal models in planning and control: evidence from grip force adjustments during movements of hand-held loads. *Journal of Neuroscience, 15,* 1519-1528.

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature, 424,* 769-771.

Fortin, C., & Couture, E. (2002).Short-term memory and time estimation: beyond the 2-second "critical" value. *Canadian Journal of Experimental Psychology, 56,* 120-127.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology, 35,* 1-36.

Furstenau, N. (2014). Simulating bistable perception with interrupted ambiguous stimulus using self-oscillator dynamics with percept choice bifurcation. *Cognitive Processing, 15,* 467-490.

Geissler, H.-G., & Kompass, R. (2001). Temporal constraints on binding? Evidence from quantal state transitions in perception. *Visual Cognition, 8,* 679-696.

Gerstner, G. E., & Cianfarani, T. (1998). Temporal dynamics of human masticatory sequences. *Physiology and Behavior, 64,* 457-461.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Current Opinion in Neurobiology, 7,* 170-184.

Gibbons, H., & Rammsayer, T. H. (2004). Current-source density analysis of slow brain potentials during time estimation. *Psychophysiology, 41,* 861-874.

Gómez, C., Argandoña, E. D., Solier, R. G., Angulo, J. C., & Vázquez, M. (1995). Timing and competition in networks representing ambiguous figures. *Brain and Cognition, 29,* 103-114.

Gori, S., Hamburger, K., & Spillmann, L. (2006). Reversal of apparent motion in the Enigma-figure with and without motion adaptation and the effect of T-junctions. *Vision Research, 46,* 3267-3273.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: the time course of natural scene understanding. *Psychological Science, 20,* 464-472.

Gregoriou, G. G., Paneri, S., & Sapountzis, P. (2015). Oscillatory synchrony as a mechanism of attentional processing. *Brain Research, 1626,* 165-182.

Groen, I. I. A., Ghebreab, S., Prins, H., Lamme, V. A. F., & Scholte, H. S. (2013). From image statistics to scene gist: evoked neural activity reveals transition from low-level natural image structure to scene category. *Journal of Neuroscience, 33,* 18814-18824.

Grondin, S., Laflamme, V., & Mioni, G. (2015). Do not count too slowly: evidence for a temporal limitation in short-term memory. *Psychonomic Bulletin and Review, 22,* 863-868.

Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: performance structures in sentence production. *Cognitive Psychology, 11,* 58-81.

Grothe, B. (2003). New roles for synaptic inhibition in sound localization. *Nature Reviews Neuroscience, 4,* 1-11.

Hanslmayr, S., Gross, J., Klimesch, W., & Shapiro, K. L. (2011). The role of alpha oscillations in temporal attention. *Brain Research Reviews, 67,* 331-343.

Hard, B. M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of Experimental Psychology: General, 140,* 586-604.

Hasson, U., Chen, J., & Honey, C. J. (2015). Hierarchical process memory: memory as an integral component of information processing. *Trends in Cognitive Sciences, 19,* 304-313.

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience, 28,* 2539-2550.

Henning, G. B., & Gaskell, H. (1981). Monaural phase sensitivity with Ronken's paradigm. *Journal of the Acoustical Society of America, 70,* 1669-1673.

Hochstein, S., & Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron, 36,* 791-804.

Hohwy, J. (2013). *The Predictive Mind.* Oxford: Oxford University Press.

Holcombe, A., & Seizova-Cajic, T. (2008). Illusory motion reversals from unambiguous motion with visual, proprioceptive, and tactile stimuli. *Vision Research, 48,* 1743-1757.

Honey, C. J., Thesen, T., Donner, T. H., Silbert, L. J., Carlson, C. E., Devinsky, O., Doyle, W. K., Rubin, N., Hegger, D. J., & Hasson, U. (2012). Slow cortical dynamics and the accumulation of information over long timescales. *Neuron, 76,* 423-434.

Ilg, R., Burazanis, S., Wohlschläger, A. M., Wöller, A., Wagenpfiel, S., & Mühlau, M. (2008). Stimulus frequency influences spontaneous perceptual reversals in ambiguous apparent motion. *Perception and Psychophysics, 70,* 437-442.

Jacob, J., Breitmeyer, B. G., & Treviño, M. (2013). Tracking the first two seconds: three stages of visual information processing? *Psychonomic Bulletin and Review, 20,* 1114-1119.

Jacoby, N., & Repp, B. H. (2012). A general linear framework for the comparison and evaluation of models of sensorimotor synchronization. *Biological Cybernetics, 106,* 135-154.

James, W. (1890). *The Principles of Psychology*. New York: Holt.

Jensen, O., Gips, B., Bergmann, T. O., & Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends in Neurosciences, 37,* 357-369.

Kagerer, F. A., Wittmann, M., Szelag, E., & von Steinbüchel, N. (2002). Cortical involvement in temporal reproduction: evidence for differential roles of the hemispheres. *Neuropsychologia, 40,* 357-366.

Kahan, T. A., & Enns, J. T. (2014). Long-term memory representations influence perception before edges are assigned to objects. *Journal of Experimental Psychology: General, 143,* 566-574.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology, 24,* 174-219.

Kanizsa, G. (1979). *Organization in vision: Essays on gestalt perception*. New York: Praeger.

Kien, J., & Kemp, A. (1994). Is speech temporally segmented? Comparison with temporal segmentation in behavior. *Brain and Language, 46,* 662-682.

Kogo, N., Hermans, L., Stuer, D., van Ee, R., & Wagemans, J. (2015). Temporal dynamics of different cases of bi-stable figure-ground perception. *Vision Research, 106,* 7-19.

Krivokapić, J. (2007). Prosodic planning: effects of phrasal length and complexity on pause duration. *Journal of Phonetics, 35,* 162-179.

Kwon, O.-S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences of America, 112,* 8142-8147.

Lee, K. (2014). Perceptual reversal guided by integration between bottom-up input and top-down feedback over time course. *Psychologia, 57,* 12 - 30.

Lehmann, D., Strik, W. K., Henggeler, B., Koenig, T., & Koukkou, M. (1998). Brain electric microstates and momentary conscious mind states as building blocks of spontaneous thinking: I. Visual imagery and abstract thoughts. *International Journal of Psychophysiology, 29,* 1-11.

Lemke, M. R., & Schleidt, M. (1999). Temporal segmentation of human short-term behavior in everyday activities and interview sessions. *Naturwissenschaften, 86,* 289-292.

Lerner, Y., Honey, C. J., Silbert, L. J., & Hasson, U. (2011). Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *Journal of Neuroscience, 31,* 2906-2915.

Lerner, Y., Honey, C. J., Katkov, M., & Hasson, U. (2014). *Journal of Neurophysiology, 111,* 2433-2444.

Lisman, J. E., & Jensen, O. (2013). The theta-gamma neural code. *Neuron, 77,* 1002-1016.

Logan, G. D. (1985). Skill and automaticity: relations, implications, and future directions. *Canadian Journal of Psychology, 39,* 367-386.

Long, G. M., Toppino, T. C., & Mondin, G. W. (1992). Prime time: fatigue and set effects in the perception of reversible figures. *Perception and Psychophysics, 52,* 609-616.

Luck, S. J., & Vogel, E. K. (2013). Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends in Cognitive Sciences, 17,* 391-400.

Macar, F., Grondin, S., & Casini, L. (1994). Controlled attention sharing influences time estimation. *Memory and Cognition, 22,* 673-686.

Mackevicius, E. L., Best, M. D., Saal, H. P., & Bensmaia, S. J. (2012). Millisecond precision spike timing shapes tactile perception. *Journal of Neuroscience, 32,* 15309-15317.

Mates, J., Radil, T., Müller, U., & Pöppel, E. (1994). Temporal integration in sensorimotor synchronization. *Journal of Cognitive Neuroscience, 6,* 332-340.

Matsuda, S., Matsumoto, H., Furubayashi, T., Hanajima, R., Tsuji, S., Ugawa, Y., & Terao, Y. (2015). The 3-second rule in hereditary pure cerebellar ataxia: a synchronized tapping study. *Plos One, 10,* e0118592.

Matthews, W. J., & Meck, W. H. (2016). Temporal cognition: connecting subjective time to perception, attention, and memory. *Psychological Bulletin, 142,* 865-907.

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision, 4,* 539-551.

Miyake, Y., Onishi, Y., & Pöppel, E. (2004). Two types of anticipation in synchronization tapping. *Acta Neurobiologiae Experimentalis, 64,* 415-426.

Montemayor, C., & Wittmann, M. (2014). The varieties of presence: hierarchical levels of temporal integration. *Timing and Time Perception, 2,* 325-338.

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin, 132,* 297-326.

Näätänen, R., Astikainen, P., Ruusvirta, T., & Huotilainen, M. (2010). Automatic auditory intelligence: an expression of the sensory-cognitive core of cognitive processes. *Brain Research Reviews, 64,* 123-136.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology, 118,* 2544-2590.

Nagy, E. (2011). Sharing the moment: the duration of embraces in humans. *Journal of Ethology, 29,* 389-393.

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature, 395,* 894-896.

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology, 28,* 28-38.

Oberauer, K., & Lin, H.-Y. (2017). An interference model of visual working memory. *Psychological Review, 124,* 21-59.

Öğmen, H., Ekiz, O, Huynh, D., Bedell, H. E., & Tripathy, S. P. (2013). Bottlenecks of motion processing during a visual glance: the leaky flask model. *Plos One, 8,* e83671.

Orbach, J., Ehrlich, D., & Heath, H. A. (1963). Reversibility of the Necker cube: I. An examination of the concept of "satiation of orientation". *Perceptual and Motor Skills, 17,* 439-458.

Orbach, J., Zucker, E., & Olson, R. (1966). Reversibility of the Necker cube: VII. Reversa rate as a function of figure-on and figure-off durations. *Perceptual and Motor Skills, 22,* 615-618.

Ossmy, O., Moran, R., Pfeffer, T., Tsetsos, K., Usher, M., & Donner, T. H. (2013). The timescale of perceptual evidence integration can be adapted to the environment. *Current Biology, 23,* 981-986.

Peters, M. (1989). The relationship between variability of intertap intervals and interval duration. *Psychological Research, 51,* 38-42.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193-198.

Pettijohn, K. A., Thompson, A. N., Tamplin, A. K., Krawietz, S. A., & Radvansky, G. A. (2016). Event boundaries and memory improvement. *Cognition, 148,* 136-144.

Po, J. M. C., Kieser, J. A., Gallo, L. M., Tésenyi, A. J., Herbison, P., & Farella, M. (2011). Time-frequency analysis of chewing activity in the natural environment. *Journal of Dental research, 90,* 1206-1210.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication, 41,* 245-255.

Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends in Cognitive Sciences, 1,* 56-61.

Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philosophical Transactions of the Royal Society: B, Biological Sciences, 364,* 1887-1896.

Radilová, J., & Pöppel, E. (1990). The perception of figure reversal as a function of contrast reversal exemplified with the Schröder staircase. *Acta Neurobiologiae Experimentalis, 50,* 37-40.

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty intervals: cognitive and perceptual factors. *Perception and Psychophysics, 50,* 565-574.

Rammsayer, T. H., & Ulrich, R. (2011). Elaborative rehearsa; of nontemporal information interferes with temporal processing of durations in the range of seconds but not milliseconds. *Acta Psychologica, 137,* 127-133.

Räsänen, O., & Laine, U. K. (2013). Time-frequency integration characteristics of hearing are optimized for perception of speech-like acoustic patterns. *Journal of the Acoustical Society of America, 134,* 407-419.

Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition, 7,* 17-42.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology, 53,* 245-277.

Repp, B. H., & Doggett, R. (2007). Tapping to a very slow beat: a comparison of musicians and non-musicians. *Music Perception, 24,* 367-376.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006-2012). *Psychonomic Bulletin and Review, 20,* 403-452.

Ritter, W., Sussman, E., Molholm, S., & Foxe, J. J. (2002). Memory reactivation or reinstatement and the mismatch negativity. *Psychophysiology, 39,* 158-165.

Roberts, B. M., Hsieh, L.-T., & Ranganath, C. (2013). Oscillatory activity during maintenance of spatial and temporal information in working memory. *Neuropsychologia, 51,* 349-357.

Roll, M., Lindgren, M., Alter, K., & Horne, M. (2012). Time-driven effects on parsing during reading. *Brain and Language, 121,* 267-272.

Roll, M., Gisselke, S., Lindgren, M., & Horne, M. (2013). Time-driven effects on processing grammatical agreement. *Frontiers in Psychology, 4,* No. 1004.

Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain: disentangling the contribution of low-level visual cues from face perception. *Vision Research, 51,* 1297-1311.

Rubin, E. (1921). *Visuell wahrgenommene figuren*. Copenhagen: Glydenalske bogahndel.

Sachs, J. S. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics, 2,* 437-442.

Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory and Cognition, 2,* 95-100.

Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: neuromagnetic evidence. *Journal of Cognitive Neuroscience, 5,* 363-370.

Samuel, A. G., & Kat, D. (2003). Inhibition of return: a graphical meta-analysis of its time course and an empirical test of its temporal and spatial properties. *Psychonomic Bulletin and Review, 10,* 897-906.

Schill, K., & Zetzsche, C. (1995). A model of visual spatio-temporal memory: the icon revisited. *Psychological Research, 57,* 88-102.

Schleidt, M. (1988). A universal time constant operating in human short-term behaviour repetitions. *Ethology, 77,* 67-75.

Schleidt, M., & Eibl-Eibesfeldt, I. (1987). A universal constant in temporal segmentation of human short-term behavior. *Naturwissenschaften, 74,* 289-290.

Schleidt, M., & Feldhütter, I. (1989). Universal time constant in human short-term behavior. *Naturwissenschaften, 76,* 127-128.

Schleidt, M., & Kien, J. (1997). Segmentation in behavior and what it can tell us about brain function. *Human Nature, 8,* 77-111.

Scotto, M. (1991). Smooth periodic eye movements can entrain perceptual alternation. *Perceptual and Motor Skills, 73,* 835-843.

Seilheimer, R. L., Rosenberg, A., & Angelaki, D. E. (2013). Models and processes of multisensory cue combination. *Current Opinion in Neurobiology, 25,* 38-46.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127-190.

Snowden, R. J., & Braddick, O. J. (1991). The temporal integration and resolution of velocity signals. *Vision Research, 31,* 907-914.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied, 74,* No. 498.

Stefanics, G., Kremláček, J., & Czigler, I. (2014). Visual mismatch negativity: a predictive coding view. *Frontiers in Human Neuroscience, 8,* No. 666.

Stephens, G. J., Honey, C. J., & Hasson, U. (2013). A place for time: the spatiotemporal structure of neural dynamics during natural audition. *Journal of Neurophysiology, 110,* 2019-2026.

Steriade, M., Nuñez, A., & Amzica, F. (1993). A novel slow oscillation (<1 Hz) of neocortical neurons in vivo: depolarizing and hyperpolarizing components. *Journal of Neuroscience, 13,* 3252-3265.

Strüber, D., & Stadler, M. (1999). Differences in top-down infliuences on the reversal rate of different categories of reversible figures. *Perception, 28,* 1185-1196.

Swallow, K. M., Zacks, J. M., & Abrams, R. A. (2009). Event boundaries in perception affect memory encoding and updating. *Journal of Experimental Psychology: General, 138,* 236-257.

Szelag, E., Kowalska, J., Galkowski, T., & Pöppel, E. (2004). Temporal processing deficits in high-functioning children with autism. *British Journal of Psychology, 95,* 269-282.

Szelag, E., Kowalska, J., Rymarczyk, K., & Pöppel, E. (2002). Duration processing in children as determined by time reproduction: implications for a few seconds temporal window. *Acta Psychologica, 110,* 1-19.

Szelag, E., von Steinbüchel, N., Reiser, M., de Langen, E. G., & Pöppel, E. (1996). *Acta Neurobiologiae Experimentalis, 56,* 215-225.

Szelag, E., von Steinbüchel, N., & Pöppel, E. (1997). Temporal processing disorders in patients with Broca's aphasia. *Neuroscience Letters, 235,* 33-36.

Takano, K., & Miyake, Y. (2007). Two types of phase correction mechanism involved in synchronized tapping. *Neuroscience Letters, 417,* 196-200.

Tanida, K., & Pöppel, E. (2006). A hierarchical model of operational anticipation windows in driving an automobile. *Cognitive Processing, 7,* 275-287.

Tapia, E., & Beck, D. M. (2014). Probing feedforward and feedback contributions yo awareness with visual masking and transcranial magnetic stimulation. *Frontiers in Psychology, 5,* No. 1173.

Toppino, T. C., & Long, G. M. (2015). Time for a change: what dominance durations reveal about adaptation effects in the perception of a bi-stable reversible figure. *Attention, Perception, and Psychophysics, 77,* 867-882.

Treisman, M. (1963). Temporal discrimination and the indifference interval. *Psychological Monographs: General and Applied, 77,* Whole No. 576.

Tsetsos, K., Pfeffer, T., Jentgens, P., & Donner, T. H. (2015). Action planning and the timescale of evidence accumulation. *Plos One, 10,* e0129473.

Ulbrich, P., Churan, J., Fink, M., & Wittmann, M. (2007). Temporal reproduction: further evidence for two processes. *Acta Psychologica, 125,* 51-65.

van der Wel, R. P. R. D., Sternad, D., & Rosenbaum, D. A. (2010). Moving the arm at different rates: slow movements are avoided. *Journal of Motor Behavior, 42,* 29-36.

VanRullen, R., Zoeful, B., & Ilhan, B. (2014). On the cyclic nature of perception in vision versus audition. *Philosophical Transactions of the Royal Society: B, Biological Sciences, 369,* 1-15.

Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia, 45,* 598-607.

Varela, F. J., (1999). Present-time consciousness. *Journal of Consciousness Studies, 6,* 111-140.

Vollrath, M., Kazenwadel, J., & Krüger, H.-P. (1992). A universal constant in temporal segmentation of human speech. *Naturwissenschaften, 79,* 479-480.

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Attention, Perception, and Psychophysics, 72,* 871-884.

Wang, L., Lin, X., Zhou, B., Pöppel, E., & Bao, Y. (2015). Subjective present: a window of temporal integration indexed by mismatch negativity. *Cognitive Processing, 16,* S131-S135.

Weber, A. I., Saal, H. P., Liber, J. D., Cheng, J.-W., Manfredi, L. R., Dammann, J. F., & Bensmaia, S. J. (2013). Spatial and temporal codes mediate the tactile perception of natural textures. *Proceedings of the National Academy of Sciences of America, 110,* 17107-17112.

Wernery, J., Atmanspacher, H., Kornmeier, J., Candia, V., Folkers, G., & Wittmann, M. (2015). Temporal processing in bistable perception of the Necker cube. *Perception, 44,* 157-168.

Wilder, D. A. (1978a). Effect of predictability on units of perception and attribution. *Personality and Social Psychology Bulletin, 4,* 281-284.

Wilder, D. A. (1978b). Predictability of behaviors, goals, and unit of perception. *Personality and Social Psychology Bulletin, 4,* 604-607.

Wing, A. M., & Kristofferson, A. B. (1973). The timing of interresponse intervals. *Perception and Psychophysics, 13,* 455-460.

Wittmann, M. (1999). Time perception and temporal processing levels of the brain. *Chronobiology International, 16,* 17-32.

Wittmann, M. (2009). The inner experience of time. *Philosophical Transactions of the Royal Society: B, Biological Sciences, 364,* 1955-1967.

Wittmann, M. (2011). Moments in time. *Frontiers in Integrative Neuroscience, 5,* No. 66.

Wittmann, M. (2013). The inner sense of time: how the brain creates a representation of duration. *Nature Reviews Neuroscience, 14,* 217-223.

Wittmann, M., & Pöppel, E. (1999-2000). Temporal mechanisms of the brain as fundamentals of communication - with special reference to music perception and performance. *Musicae Scientiae, 3 (SI),* 13-28.

Woodrow, H. (1930). The reproduction of temporal intervals. *Journal of Experimental Psychology, 13,* 473-499.

Woodrow, H. (1932). The effect of rate of sequence upon the accuracy of synchronization. *Journal of Experimental Psychology, 15,* 357-379.

Wutz, A., & Melcher, D. (2014). The temporal window of individuation limits visual capacity. *Frontiers in Psychology, 5,* No. 952.

Yoshimoto, S., Uchida-Ota, M., & Takeuchi, T. (2014). The reference frame of visual motion priming depends on underlying motion mechanisms. *Journal of Vision, 14* (1), No. 10.

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28,* 979-1008.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130,* 29-58.

Zimmermann, E., Morrone, M. C., & Burr, D. C. (2014). Buildup of spatial information over time and across eye-movements. *Behavioural Brain Research, 275,* 281-287.

Table 1

Mean duration of segments constructed under instruction to segment behaviour sequences into the smallest units that seem natural

Duration	Study	Notes
5.8 s	Newtson (1973)	Solo actor, 5 min scripted behaviour sequence.
3.5 s	Wilder (1978a)	Solo actor, predictable behaviour sequence.
4.2 s	Wilder (1978a)	Solo actor, unpredictable behaviour sequence.
8.9 s	Wilder (1978b)	Solo actor, predictable behaviour sequence.
5.4 s	Wilder (1978b)	Solo actor, unpredictable behaviour sequence.
12.8 s	Zacks et al. (2001)	Two actors, various activities, scripted behaviour.
6.29 s	Zacks (2004, expt. 1)	Animated moving shapes, "random" instructions.[1]
9.95 s	Zacks (2004, expt. 1)	Animated moving shapes, "intentional" instructions.[2]
15.2 s	Zacks (2004, expt. 2)	Animated moving shapes, "random" instructions.[1]
14.7 s	Zacks (2004, expt. 2)	Animated moving shapes, "intentional" instructions.[2]
13.8 s	Zacks (2004, expt. 3)	Animated moving shapes, "random" instructions.[1]
12.2 s	Zacks (2004, expt. 3)	Animated moving shapes, "intentional" instructions.[2]
15.1 s	Zacks (2004, expt. 3)	Animated moving shapes, random/game.[3]
13.2 s	Zacks (2004, expt. 3)	Animated moving shapes, intentional/game.[3]
19.1 s[4]	Swallow et al. (2009)	Movie clips, multiple actors, everyday activities.
6.8 s	Hard et al. (2011)	Solo actor, everyday activities, slideshow.[5]
10.7 s	Kurby & Zacks (2011)	Solo actors, everyday events, young participants.

15.3 s	Kurby & Zacks (2011)	Solo actors, everyday events, old participants.

Notes. 1. The materials comprised an animation of two moving geometrical objects, moving "randomly with momentum approximating random forces in the presence of friction" (Zacks, 2004, p. 984), 600 s in duration. Participants were instructed that the movements were random. 2. As note 1, but participants were instructed that the animation "depicted the recorded movements of two people in a room completing a goal-directed activity" (Zacks, 2004, p. 985). 3. In Zacks (2004, experiment 3), the animation was constructed by two actors using a keyboard and generating movement patterns that corresponded to several specific kinds of interactions, such as chasing. In the "game" condition, participants were informed about how the animations were constructed. 4. Mean of means for six different movie clips. Range of means was 17.5 to 22.0 s. 5. Activities were filmed, but stimuli were slides sampled from the recording at 1 s intervals.

Table 2

Mean reversal intervals for the Necker cube

Reversal rate	Study
4.6 s	Orbach, Ehrlich, & Heath (1963)
7.5 s	Long, Toppino, & Mondin (1992)
6.6 s	Gomez, Argandoña, Solier, Angulo, & Váquez (1995)
5.2 s	Meng & Tong (2004)
4.5 s	Sauer, Lemke, Wittmann, Kohls, Mochty, & Walach (2012)
4.4 s	Kogo, Hermans, Stuer, van Ee, & Wagemans (2015)